

DÉPARTEMENT DES AFFAIRES
ÉCONOMIQUES ET SOCIALES

Division de statistique

Études méthodologiques

Série F n° 96

Enquêtes sur les ménages dans les pays en développement et les pays en transition



Nations Unies
New York, 2007

Département des affaires économiques et sociales

Le Département des affaires économiques et sociales du Secrétariat des Nations Unies est un intermédiaire essentiel entre les politiques économiques, sociales et environnementales à l'échelle mondiale et l'action nationale. Ce département exerce ses activités dans trois domaines interdépendants : i) il recueille, établit et analyse un large éventail de données et informations économiques, sociales et environnementales dans lesquelles les États membres des Nations Unies puisent pour examiner des problèmes courants et s'inspirer des possibilités d'action; ii) il facilite les négociations entre États membres au sein de nombreux organes intergouvernementaux sur le choix d'actions communes à mener pour faire face à l'émergence de problèmes mondiaux; et iii) il conseille les gouvernements intéressés sur les moyens de passer des cadres d'action élaborés dans le cadre des conférences et des sommets des Nations Unies à des programmes à l'échelon national et, par le biais de l'assistance technique, il aide à renforcer les capacités nationales..

Note

Les cotes des documents de l'Organisation des Nations Unies se composent de lettres majuscules et de chiffres. La simple mention d'une cote dans un texte signifie qu'il s'agit d'un document de l'Organisation.

Préface

Les enquêtes sur les ménages sont une importante source de données socio-économiques. Elles sont souvent la source d'importants indicateurs qui permettent d'étayer et de suivre les politiques de développement. Dans les pays en développement, elles sont à présent l'un des principaux moyens de recueillir des données, complétant et parfois remplaçant d'autres programmes de collecte de données et systèmes de registre d'état civil.

La présente publication décrit les dernières réalisations concernant plusieurs aspects importants de l'exécution d'enquêtes sur les ménages dans les pays en développement et les pays en transition, y compris la conception d'échantillons, la mise en œuvre des enquêtes, les erreurs non dues à l'échantillonnage, les coûts des enquêtes et l'analyse des données d'enquête. Ce manuel a pour principal but d'aider les statisticiens chargés d'enquêtes nationales à concevoir les enquêtes sur les ménages d'une manière efficace et fiable et de permettre aux usagers d'utiliser davantage les données produites par ces enquêtes.

Les 25 chapitres de cette publication sont l'œuvre d'éminents experts mondiaux des méthodes de recherche par enquête. La plupart de ces experts ont une expérience pratique de l'assistance aux responsables nationaux des statistiques des pays en développement et des pays en transition. Cet ouvrage présente certains traits uniques :

- Un examen particulier des besoins des pays en développement et des pays en transition;
- L'attention portée aux normes et aux caractéristiques opérationnelles applicables à divers pays et à diverses enquêtes;
- Le traitement des coûts des enquêtes, avec des exemples concrets de l'établissement de leur budget et des analyses de leurs coûts par composante;
- Un traitement exhaustif des erreurs non dues à l'échantillonnage;
- L'examen des techniques élémentaires et perfectionnées de l'analyse des données des enquêtes sur les ménages, y compris une comparaison empirique des derniers logiciels conçus pour l'analyse de données d'enquêtes complexes;
- La présentation d'exemples de conception, de mise en œuvre et d'analyse des données de certaines enquêtes sur les ménages effectuées dans les pays en développement et les pays en transition;
- La présentation de plusieurs études de cas portant sur de véritables enquêtes à grande échelle effectuées dans des pays en développement et des pays en transition, qui peuvent servir d'exemples à suivre pour la conception d'enquêtes similaires.

Cette publication se fonde sur de précédentes initiatives de la Division de statistique du Département des affaires économiques et sociales (DESA/UNSD) pour améliorer la qualité de la méthodologie des enquêtes et renforcer la capacité des systèmes nationaux de statistique. La plus complète de ces initiatives des deux dernières décennies est le Programme de mise en place de dispositifs nationaux d'enquête sur

les ménages (NHSCP). Ce programme avait pour but d'aider les pays en développement à recueillir des données démographiques et socio-économiques critiques au moyen d'un système intégré d'enquêtes sur les ménages afin d'aider à la planification du développement, à la formulation de politiques et à la mise en œuvre de programmes. Il a beaucoup contribué au développement des statistiques dans de nombreux pays en développement, en particulier en Afrique, qui a lui-même bénéficié d'un fort accroissement du nombre et des types d'enquêtes effectuées dans les années 1980. Par ailleurs, le NHSCP a favorisé un travail méthodologique qui a permis la publication de plusieurs manuels et études techniques. Le *Manuel des méthodes d'enquêtes sur les ménages* (édition révisée)¹ donne un aperçu général des questions liées à la conception et la mise en œuvre des enquêtes sur les ménages. Il a été suivi d'une série de publications traitant de certaines questions et procédures relatives à des domaines spécifiques des méthodes d'enquête et portant sur plusieurs domaines, parmi lesquels :

¹ *Études méthodologiques*, n° 31
(publication des Nations Unies,
numéro de vente : E.83.XVII.13).

- *Programme de mise en place de dispositifs nationaux d'enquête sur les ménages : Sampling Frames and Sample Designs for Integrated Household Survey Programmes, Preliminary Version* (DP/UN/INT-84-014/5E), New York, 1986
- *Programme de mise en place de dispositifs nationaux d'enquête sur les ménages : erreurs d'échantillonnage dans les enquêtes sur les ménages* (UNFPA/UN/INT-92-P80-15E), New York, 1993
- *Programme de mise en place de dispositifs nationaux d'enquête sur les ménages : traitement des données d'enquête : problèmes et procédures* (DP/UN/INT-81-041/1), New York, 1982
- *Programme de mise en place de dispositifs nationaux d'enquête sur les ménages : erreurs autres que l'échantillonnage dans les enquêtes sur les ménages : sources, évaluation et maîtrise : version préliminaire* (DP/UN/INT-81-041/2), New York, 1982
- *Programme de mise en place de dispositifs nationaux d'enquête sur les ménages : élaboration et conception des questionnaires d'enquête* (INT-84-014), New York, 1985
- *Programme de mise en place de dispositifs nationaux d'enquête sur les ménages : Household Income and Expenditure Surveys: A Technical Study* (DP/UN/INT-88-X01/6E), New York, 1989
- *Programme de mise en place de dispositifs nationaux d'enquête sur les ménages : Guidelines for Household Surveys on Health* (INT/89/X06), New York, 1995
- *Programme de mise en place de dispositifs nationaux d'enquête sur les ménages : Sampling Rare and Elusive Populations* (INT-92-P80-16E), New York, 1993

La présente publication actualise et développe les aspects techniques des problèmes et procédures traités de façon détaillée dans les publications susmentionnées, tout en se concentrant exclusivement sur leur application aux enquêtes dans les pays en développement et les pays en transition.

Paul CHEUNG
Directeur,
Division de statistique des Nations Unies,
Département des affaires économiques et sociales

Aperçu général

Cette publication s'articule en deux parties, formant un total de 25 chapitres. La première partie est composée de 21 chapitres, regroupés en cinq sections A à E. Voici un résumé du contenu de chaque section de cette première partie.

Section A : *Conception et mise en œuvre des enquêtes*. Cette section comprend trois chapitres. Le chapitre II présente un aperçu des divers problèmes qui se posent lors de la conception des enquêtes sur les ménages dans les pays en développement et les pays en transition. Les chapitres III et IV examinent respectivement les problèmes concernant la conception des questionnaires et les problèmes liés à l'exécution des enquêtes dans les pays en développement et les pays en transition.

Section B : *Conception de l'échantillon*. Cette section contient une note d'introduction et trois chapitres qui traitent des aspects particuliers de la conception des échantillons. Le chapitre V traite de la conception des échantillons maîtres et des cadres directeurs. L'utilisation d'effets liés à la conception dans la conception et l'analyse d'un échantillon est examinée au chapitre VI, et le chapitre VII offre une analyse empirique des effets liés à la conception dans les enquêtes effectuées dans plusieurs pays en développement.

Section C : *Erreurs autres que les erreurs d'échantillonnage*. Cette section contient une note d'introduction et quatre chapitres qui traitent de divers aspects de la mesure, de l'évaluation et de la maîtrise des erreurs autres que les erreurs d'échantillonnage dans les pays en développement et les pays en transition. Le chapitre VIII traite de l'erreur de non-observation (non-réponse et non-couverture). Les erreurs de mesure sont traitées au chapitre IX. Le chapitre X présente des directives et procédures d'assurance de la qualité applicables aux Enquêtes mondiales sur la santé, programme d'enquêtes dans les pays en développement, parrainées par l'Organisation mondiale de la santé (OMS). Le chapitre XI décrit une étude de cas sur la mesure, l'évaluation et la compensation des erreurs autres que les erreurs d'échantillonnage dans les enquêtes sur les ménages au Brésil.

Section D : *Coûts des enquêtes*. Cette section contient une note d'introduction et trois chapitres. Le chapitre XII offre un cadre général pour l'analyse des coûts des enquêtes effectuées dans les pays en développement et les pays en transition. À partir de données empiriques, le chapitre XIII décrit un modèle de coûts pour une enquête sur les revenus et les dépenses effectuée dans un pays en développement. Le chapitre XIV traite des questions relatives à l'élaboration d'un budget pour les myriades de phases et de fonctions d'une enquête sur les ménages et présente un certain nombre d'exemples et d'études de cas qui sont utilisés pour établir des comparaisons et illustrer les importantes questions budgétaires examinées dans le chapitre.

Section E : *Analyse des données de l'enquête*. Cette section contient une note d'introduction et sept chapitres consacrés à l'analyse des données d'enquête. Le chapitre XV offre des directives détaillées pour la gestion des données d'enquêtes sur les ménages. Le chapitre XVI examine l'analyse tabulaire de base des données d'enquête et présente plusieurs exemples concrets. Le chapitre XVII porte sur l'utilisation d'enquêtes à thèmes multiples ou polyvalentes sur les ménages en tant qu'instrument de lutte contre la pauvreté dans les pays en développement. Le chapitre XVIII traite de l'utilisa-

tion de méthodes statistiques multivariates pour l'établissement d'indices à partir des données d'enquêtes sur les ménages. Le chapitre XIX a trait à l'analyse statistique des données d'enquêtes, et se concentre sur les techniques de base d'analyse fondées sur des modèles, à savoir la régression linéaire multiple, la régression logistique et les méthodes multiniveaux. Le chapitre XX présente des approches plus avancées de l'analyse des données d'enquête qui tiennent compte des effets de la complexité de la conception sur l'analyse. Enfin, le chapitre XXI examine les diverses méthodes utilisées pour l'estimation des erreurs d'échantillonnage dans les données d'enquête et décrit des techniques d'analyse pratique des données, comparant plusieurs logiciels utilisés pour analyser les données d'enquêtes complexes. Il fait également ressortir la relation étroite qui existe entre la conception de l'échantillon et l'analyse des données. On trouvera plus de détails sur la comparaison des logiciels, notamment sur les résultats informatiques de divers logiciels, dans le CD-ROM qui accompagne la présente publication.

La seconde partie de cette publication, qui contient quatre chapitres précédés d'une note d'introduction, est consacrée à des études de cas qui offrent des exemples concrets d'enquêtes effectuées dans des pays en développement et des pays en transition. Ces chapitres présentent un examen détaillé et systématique d'enquêtes payées par les usagers et parrainées par des institutions internationales et d'enquêtes financées sur le budget des pays et effectuées dans le cadre des programmes courants d'enquêtes des systèmes statistiques nationaux. Le programme d'Enquêtes démographiques et sanitaires (DHS) est décrit au chapitre XXII, et le programme d'étude de la mesure des niveaux de vie (LSMS) est décrit au chapitre XXIII. L'examen de ces deux séries d'enquêtes comprend le calcul des effets liés à la conception des estimations d'un certain nombre de caractéristiques essentielles. Le chapitre XXIV traite de la conception et de la mise en œuvre des enquêtes sur les ménages en utilisant comme illustration une enquête effectuée en République populaire démocratique lao. Le chapitre XXV est consacré aux aspects généraux de la conception et de la mise en œuvre d'enquêtes effectuées dans les pays en transition, et comprend plusieurs études de cas.

Remerciements

La préparation d'une publication de cette ampleur doit nécessairement être un travail de coopération. DESA/UNSD a tiré un immense profit du précieux concours de nombreux consultants individuels et d'organisations à travers le monde, tant internes qu'externes au système des Nations Unies. Ces consultants sont des experts qui possèdent une solide expérience de la conception, la mise en œuvre et l'analyse d'enquêtes complexes, bien souvent dans des pays en développement ou des pays en transition.

Tous les chapitres de cette publication ont fait l'objet d'examens très rigoureux par des pairs. Tout d'abord, chaque chapitre a été examiné par deux éditeurs réputés pour leur connaissance du domaine traité. Les chapitres ainsi révisés ont été réunis pour former la première version de cette publication, qui a elle-même fait l'objet d'un examen critique, lors de la réunion du groupe d'experts organisée par DESA/UNSD à New York en octobre 2002. À l'issue de cette réunion, un comité d'édition a été établi pour revoir la publication et faire les dernières recommandations concernant sa structure et sa teneur. Cette phase du processus d'examen a abouti à une restructuration et à un allègement de l'ensemble de la publication, destinés à la rendre plus cohérente, plus complète et plus homogène. De nouveaux chapitres ont été ajoutés et d'autres ont été révisés selon les recommandations de la réunion du groupe d'experts et du comité de révision. Chaque chapitre ainsi révisé a ensuite fait l'objet d'un troisième examen par deux éditeurs avant que soit prise la décision finale de l'incorporer ou non à la publication. Une équipe d'éditeurs a ensuite procédé à un examen final de la publication dans son intégralité, afin de s'assurer que l'ouvrage présenté était techniquement sain, homogène et fidèle aux objectifs initiaux de la publication.

DESA/UNSD remercie vivement M. Graham Kalton de sa contribution inestimable à cette publication. M. Graham Kalton a présidé à la fois la réunion du groupe d'experts et le comité d'édition, révisé de nombreux chapitres et offert des conseils techniques et une orientation intellectuelle au personnel de DESA/UNSD durant toute l'exécution de ce projet. M. John Eltinge a prodigué de précieux conseils aux premiers stades de l'élaboration des idées qui ont abouti à cette publication et, en qualité d'éditeur de plusieurs chapitres et d'animateur et de collaborateur pour certains des travaux de recherche générale qui ont débouché sur l'élaboration de son cadre, et a continué de jouer un rôle déterminant vis-à-vis de tous les aspects de ce projet. MM. James Lepkowski, Oladejo Ajayi, Hans Pettersson, Karol Krotki et Anthony Turner ont dispensé une aide cruciale lors de l'édition de plusieurs chapitres, ainsi que des conseils d'ordre général et un appui à divers stades du projet.

De nombreux autres experts ont collaboré à ce projet, en qualité d'auteurs de certains chapitres, d'éditeurs de chapitres rédigés par d'autres experts, ou en tant qu'auteurs et éditeurs. D'autres ont apporté leur contribution en participant à la réunion du groupe d'experts et en offrant des examens constructifs de tous les aspects de la version initiale de la publication. Les noms et affiliations de tous les experts ayant participé à ce projet figurent dans une liste après la table des matières.

Il eût été difficile, sinon impossible, d'atteindre les objectifs ambitieux de ce projet sans les immenses contributions de plusieurs membres du personnel de DESA/

UNSD à chaque stade de sa réalisation. M. Ibrahim Yansaneh a conçu la proposition relative à cette publication, recruté les autres participants et coordonné tous les aspects techniques du projet, y compris son processus d'édition. Il est également l'auteur de plusieurs chapitres et a fait fonction de rédacteur en chef de l'ensemble de la publication. Le Directeur et le Directeur-adjoint de DESA/UNSD ont prodigué des encouragements et un appui institutionnel à tous les stades du projet. M. Stefan Schweinfest a assuré la gestion de tous les aspects de la publication. Mme Sabine Warschburger a conçu et entretenu le site Internet du projet, et Mme Denise Quiroga a assuré un service de secrétariat exemplaire en facilitant la circulation des nombreux documents entre leurs auteurs et leurs éditeurs, en organisant et en harmonisant les formats et les styles d'écriture disparates de ces documents, et en aidant à faire respecter le calendrier d'exécution du projet.

TABLE DES MATIÈRES

Préface.....	iii
Aperçu général.....	vi
Remerciements.....	viii
Liste des experts ayant contribué à la réalisation de cette publication.....	xxxii
Auteurs.....	xxxiii
Éditeurs.....	xxxiv
PREMIÈRE PARTIE. Conception, mise en œuvre et analyse des enquêtes.....	1
Chapitre premier. Introduction.....	3
A. Enquêtes sur les ménages dans les pays en développement et les pays en transition.....	3
B. Objectifs de la présente publication.....	4
C. Importance pratique des objectifs.....	5
Section A. Conception et mise en œuvre des enquêtes.....	7
Chapitre II. Aperçu des problèmes de conception d'échantillons pour les enquêtes sur les ménages dans les pays en développement et les pays en transition.....	9
A. Introduction.....	9
1. Modèles d'échantillon pour les enquêtes dans les pays en développement et les pays en transition.....	9
2. Aperçu général.....	10
B. Échantillonnage stratifié à phases multiples.....	11
1. Stratification explicite.....	11
2. Stratification implicite.....	12
3. Choix des UPE de l'échantillon.....	12
4. Échantillonnage des UPE selon une probabilité proportionnelle à leur taille.....	14
5. Sélection d'un échantillon de ménages.....	16
6. Nombre de ménages à choisir par UPE.....	17
C. Cadres d'échantillonnage.....	19
1. Caractéristiques des cadres d'échantillonnage retenus pour les enquêtes dans les pays en développement et les pays en transition.....	19
2. Problèmes liés aux cadres d'échantillonnage et solutions possibles.....	20
3. Maintien et évaluation des cadres d'échantillonnage.....	21
D. Estimation des domaines.....	22
1. Nécessité d'estimations de domaines.....	22
2. Allocation d'échantillons.....	22
E. Taille des échantillons.....	23
1. Facteurs qui influent sur les décisions concernant la taille des échantillons.....	23
2. Précision des estimations de l'enquête.....	23
3. Qualité des données.....	26

4. Coût et opportunité.....	26
F. Analyse de l'enquête.....	27
1. Établissement et ajustement des pondérations d'échantillonnage.....	27
2. Analyse des données d'enquêtes sur les ménages.....	28
G. Conclusion.....	29
Annexe : Diagramme du processus d'enquête.....	32
Chapitre III. Aperçu de la conception d'un questionnaire pour les enquêtes sur les ménages dans les pays en développement.....	33
A. Introduction.....	33
B. Vue d'ensemble.....	34
1. Objectifs de l'enquête.....	34
2. Contraintes.....	36
3. Quelques conseils pratiques.....	37
C. Les détails.....	37
1. L'approche par modules.....	37
2. Formatage et homogénéité.....	39
3. Autres conseils sur certains aspects de la conception du questionnaire.....	43
D. Le processus.....	43
1. Constitution d'une équipe.....	44
2. Élaboration de la première version du questionnaire.....	44
3. Essais sur le terrain et établissement de la version finale du questionnaire.....	45
E. Conclusion.....	47
Chapitre IV. Vue d'ensemble de la mise en œuvre des enquêtes sur les ménages dans les pays en développement.....	49
A. Introduction.....	49
B. Tâches à accomplir avant que l'équipe de l'enquête se rende sur le terrain.....	50
1. Financement du budget.....	50
2. Plan de travail.....	51
3. Sélection d'un échantillon de ménages.....	53
4. Rédaction des manuels de formation.....	55
5. Formation du personnel de terrain et des agents chargés de l'entrée des données.....	56
6. Travail de terrain et plan d'entrée des données.....	56
7. Exécution d'un essai pilote.....	57
8. Lancement d'une campagne de publicité.....	57
C. Activités de soutien au travail de terrain.....	58
1. Communications et transports.....	58
2. Supervision et assurance de la qualité.....	59
3. Gestion des données.....	59
D. Activités nécessaires une fois que le travail de terrain, l'entrée et le traitement des données sont terminés.....	60
1. Compte rendu oral de mission.....	60

2. Préparation de l'ensemble des données et de la documentation finale.....	60
3. Analyse des données	61
E. Conclusion	62
Section B. Conception de l'échantillon	63
Introduction.....	65
 Chapitre V. Conception de cadres directeurs d'échantillonnage et d'échantillons- maîtres pour les enquêtes sur les ménages dans les pays en développement.....	 69
A. Introduction.....	69
B. Cadres directeurs d'échantillonnage et échantillons-maîtres : aperçu général.....	70
1. Cadres directeurs d'échantillonnage.....	70
2. Échantillons-maîtres	71
3. Résumé et conclusion	74
C. Conception d'un cadre directeur d'échantillonnage.....	75
1. Évaluation de la qualité des données et documents.....	75
2. Décision relative à la couverture du cadre directeur d'échantillonnage	76
3. Décision concernant les unités de base du cadre d'échantillonnage	77
4. Informations sur les unités du cadre d'échantillonnage à inclure dans ce cadre.....	78
5. Documentation et tenue d'un cadre directeur d'échantillonnage.....	80
D. Conception d'échantillons-maîtres	82
1. Choix des unités primaires d'échantillonnage à incorporer à l'échantillon-maître.....	82
2. Fusion/fractionnement des zones pour réduire les différences de taille entre UPE	83
3. Stratification des UPE et répartition de l'échantillon-maître entre les strates.....	85
4. Sélection de l'échantillon d'UPE.....	86
5. Durabilité des échantillons-maîtres	87
6. Documentation.....	88
7. Utilisation d'un échantillon-maître pour des enquêtes sur des entités économiques	89
E. Conclusion	90
 Chapitre VI. Estimation des composantes des effets liés à la conception à utiliser lors de la conception d'échantillons.....	 93
A. Introduction.....	93
B. Composantes des effets liés à la conception	97
1. Stratification.....	97
2. Mise en grappes.....	102
3. Ajustements des pondérations.....	105
C. Modèles pour les effets liés à la conception.....	108

D.	Utilisation des effets liés à la conception pour la conception des échantillons	112
E.	Conclusion	115
Chapitre VII.	Analyse des effets liés à la conception pour les enquêtes dans les pays en développement.....	119
A.	Introduction.....	119
B.	Les enquêtes.....	120
C.	Effets liés à la conception	122
D.	Calcul des taux d'homogénéité	128
E.	Discussion	131
Annexe :	Description des échantillons conçus pour les 11 enquêtes sur les ménages	132
Section C.	Erreurs autres que les erreurs d'échantillonnage.....	137
Introduction.....		139
Chapitre VIII.	L'erreur de non-observation dans les enquêtes sur les ménages dans les pays en développement.....	141
A.	Introduction.....	141
B.	Cadre pour la compréhension de l'erreur de non-couverture et de l'erreur de non-réponse	142
C.	Erreur de non-couverture.....	145
1.	Sources de la non-couverture	145
2.	Erreur de non-couverture.....	148
D.	Erreur de non-réponse.....	152
1.	Sources de non-réponse dans les enquêtes sur les ménages	152
2.	Distorsion due à la non-réponse.....	153
3.	Mesure de la distorsion due à la non-réponse	155
4.	Réduction et compensation des cas de non-réponse unitaire dans les enquêtes sur les ménages.....	156
5.	Non-réponse ponctuelle et imputation.....	159
Chapitre IX.	L'erreur de mesure dans les enquêtes sur les ménages : sources et évaluation	163
A.	Introduction.....	163
B.	Les sources de l'erreur de mesure	165
1.	Les effets du questionnaire.....	166
2.	Les effets liés au mode de collecte des données	168
3.	Les effets liés à l'enquêteur	171
4.	Les effets liés au répondant	173
C.	Approches de l'évaluation quantitative de l'erreur de mesure	175
1.	Expériences effectuées au hasard.....	175
2.	Méthodes de recherche cognitive.....	176
3.	Études fondées sur une seconde entrevue.....	177
4.	Études consacrées à la vérification des résultats.....	180
5.	Études de la variance liée à l'enquêteur.....	182
6.	Évaluation codée du comportement.....	182
D.	Conclusions : l'erreur de mesure	183

Chapitre X. L'assurance de qualité dans les enquêtes :	
normes, directives et procédures.....	189
A. Introduction.....	190
B. Normes et procédures d'assurance de qualité.....	190
C. Application pratique des directives d'assurance de qualité :	
l'exemple des Enquêtes sur la santé dans le monde.....	191
1. Sélection des organismes d'enquête.....	193
2. Échantillonnage.....	194
3. Traduction.....	198
D. Formation.....	200
E. Exécution de l'enquête.....	203
F. Entrée des données.....	207
G. Analyse des données.....	210
H. Indicateurs de qualité.....	210
1. Indice de déviation de l'échantillon.....	211
2. Taux de réponse.....	211
3. Pourcentage de données manquantes.....	212
4. Taux de concordance entre la première et la seconde entrevue.....	212
I. Les rapports des pays.....	213
J. Visites sur place.....	214
K. Conclusions.....	215
Chapitre XI. Établissement de rapports et compensation des erreurs	
autres que les erreurs d'échantillonnage au Brésil :	
pratique actuelle et défis pour l'avenir.....	219
A. Introduction.....	219
B. Pratique actuelle de la notification et de la compensation	
des erreurs autres que les erreurs d'échantillonnage	
dans les enquêtes sur les ménages au Brésil.....	223
1. Erreurs de couverture.....	223
2. Non-réponse.....	226
3. Erreurs de mesure et de traitement.....	230
C. Défis et perspectives.....	232
D. Autres lectures recommandées.....	234
Section D. Coûts des enquêtes.....	237
Introduction.....	239
Chapitre XII. Analyse des questions de coût concernant les enquêtes effectuées	
dans les pays en développement et les pays en transition.....	241
A. Introduction.....	241
1. Critères d'efficacité pour la conception d'échantillons.....	241
2. Éléments de la structure des coûts des enquêtes dans les pays	
en développement et les pays en transition.....	242
3. Aperçu général de ce chapitre.....	243
B. Composantes du coût des enquêtes.....	244
C. Coûts des enquêtes effectuées dans les pays	
disposant d'une vaste infrastructure.....	245

1.	Facteurs liés aux activités préparatoires	245
2.	Facteurs liés à la collecte et au traitement des données.....	245
D.	Coûts des enquêtes dans les pays où l'infrastructure nécessaire est insuffisante ou inexistante.....	246
E.	Facteurs liés aux modifications des objectifs de l'enquête.....	247
F.	Quelques mises en garde concernant la détermination des coûts des enquêtes.....	248
G.	Résumé et conclusions	249
	Annexe : Cadre budgétaire pour les enquêtes en grappes à indicateurs multiples (MICS) du Fonds des Nations Unies pour l'enfance (UNICEF)	252
Chapitre XIII.	Modèle de détermination des coûts d'une enquête sur les revenus et les dépenses	255
A.	Introduction.....	255
B.	Modèles de détermination des coûts et estimations des coûts.....	256
C.	Modèles de détermination des coûts pour une conception efficace d'échantillons.....	257
D.	Étude de cas : l'Enquête de 2002 sur les dépenses et la consommation en République démocratique populaire lao.....	259
E.	Modèle de détermination des coûts pour le travail de terrain lors de l'Enquête de 2002 sur les dépenses et la consommation (LECS-3).....	261
F.	Conclusions	263
Chapitre XIV.	Élaboration d'un cadre pour l'établissement du budget des enquêtes sur les ménages dans les pays en développement	265
A.	Introduction.....	265
B.	Considérations préliminaires	266
1.	Les phases de l'enquête.....	266
2.	Calendrier de l'enquête	267
3.	Type d'enquête.....	268
4.	Budgets et dépenses	269
5.	Études précédentes.....	269
C.	Principaux postes comptables du cadre budgétaire.....	270
1.	Personnel.....	270
2.	Transports.....	272
3.	Équipement	272
4.	Fournitures.....	272
5.	Autres dépenses	272
6.	Exemples d'établissement de budget par postes comptables.....	273
D.	Principales activités de l'enquête menées dans les limites du cadre budgétaire.....	275
1.	Établissement du budget de réparation de l'enquête	275
2.	Établissement du budget d'exécution de l'enquête	275
3.	Établissement du budget de traitement des données de l'enquête....	276
4.	Établissement du budget de préparation du rapport	276
5.	Exemples d'établissement de budget pour les activités de l'enquête	276
E.	Synthèse.....	278

F.	Limitations potentielles du budget et pièges	279
G.	Tenue de rapports et résumés	280
H.	Conclusions	281
Annexe.	Exemples de formulaires pour la tenue de listes quotidiennes et hebdomadaires d'activités.....	282
Section E. Analyse des données de l'enquête.....		285
Introduction		287
Chapitre XV.	Guide pour la gestion des données d'enquêtes sur les ménages	291
A.	Introduction.....	291
B.	Gestion des données et conception du questionnaire.....	292
C.	Stratégies opérationnelles pour l'entrée et l'édition des données.....	294
D.	Critères de contrôle de qualité.....	297
E.	Conception du programme d'entrée des données.....	299
F.	Organisation et diffusion des ensembles de données de l'enquête.....	301
G..	Gestion des données durant le processus d'échantillonnage	304
H.	Résumé des recommandations.....	316
Chapitre XVI.	Présentation de statistiques descriptives simples extraites des données d'une enquête sur les ménages.....	319
A.	Introduction.....	319
B.	Variables et statistiques descriptives.....	320
1.	Types de variables.....	320
2.	Statistiques descriptives simples.....	321
3.	Présentation de statistiques descriptives pour une variable	323
4.	Présentation de statistiques descriptives de deux variables.....	327
5.	Présentation de statistiques descriptives pour trois variables ou plus.....	330
C.	Conseils d'ordre général pour la présentation de statistiques descriptives.....	330
1.	Préparation de données	330
2.	Présentation de résultats.....	332
3.	Qu'est-ce qui constitue un bon tableau ?	333
4.	Utilisation de pondérations	335
D.	Préparation d'un rapport général (résumé) sur une enquête sur les ménages.....	336
1.	Contenu	336
2.	Processus	336
E.	Conclusions	337
Chapitre XVII.	Utilisation d'enquêtes sur les ménages à thèmes multiples pour améliorer les politiques de lutte contre la pauvreté dans les pays en développement.....	339
A.	Introduction.....	339
B.	Analyse descriptives.....	340
1.	Définir la pauvreté.....	341
2.	Construction d'un profil de pauvreté	342

3.	Utilisation des profils de pauvreté pour une simple analyse des politiques	343
C.	Analyse de régression multiple des données d'enquêtes sur les ménages	345
1.	Analyse de la demande	346
2.	Utilisation des services sociaux	347
3.	Impact de certains programmes gouvernementaux	347
D.	Résumé et conclusions	348
Chapitre XVIII. Méthodes multivariates pour la construction d'indices.....		351
A.	Introduction.....	352
B.	Quelques restrictions quant à l'utilisation des méthodes multivariates....	352
C.	Aperçu des méthodes multivariates.....	353
D.	Graphiques et mesures sommaires	354
E.	Analyse par grappes.....	356
F.	Analyse en composantes principales (ACP).....	360
G.	Méthodes multivariates de construction d'indices.....	363
1.	Modélisation des dépenses de consommation pour la construction d'un indice de substitution du revenu	364
2.	Analyse en composantes principales (ACP) utilisée pour établir un indice de « richesse »	366
H.	Conclusions	367
Chapitre XIX. Analyse statistique des données d'enquête.....		371
A.	Introduction.....	371
B.	Statistiques descriptives : pondérations et estimation de la variance.....	373
C.	Statistiques analytiques	378
D.	Commentaires généraux sur la modélisation par régression	379
E.	Modèles de régression linéaire	381
F.	Modèles de régression logistique	386
G.	Utilisation de modèles à niveaux multiples	388
H.	Modélisation à l'appui des processus d'enquête	393
I.	Conclusions	393
Chapitre XX. Approches plus avancées de l'analyse des données d'enquêtes		397
A.	Introduction.....	397
1.	Conception de l'échantillon et analyse des données	397
2.	Exemples des effets (et de l'absence d'effets) de la conception de l'échantillon sur l'analyse	398
3.	Concepts de base	400
4.	Effets de conception et leur rôle dans l'analyse des données d'un échantillon complexe	401
B.	Approches fondamentales de l'analyse des données d'échantillons complexes	402
1.	Utilisation des spécifications du modèle comme base de l'analyse...	402
2.	Relations possibles entre le modèle et la conception de l'échantillon : conceptions informatives et non informatives	403
3.	Problèmes liés à l'utilisation de logiciels courants pour l'analyse d'échantillons complexes.....	404

C.	Analyse de régression et modèles linéaires.....	405
1.	Effets des variables de conception ne figurant pas dans le modèle et des estimations de régression pondérée.....	405
2.	Évaluation de l'effet de la conception sur l'analyse de régression.....	407
3.	Modèles à niveaux multiples et échantillon de conception informative	408
D.	Analyse des données catégoriques.....	409
1.	Modifications des tests de chi-carré en tests mesurant la bonne adaptation et l'indépendance.....	409
2.	Généralisations pour les modèles log-linéaires.....	412
E.	Résumé et conclusions	413
Annexe.	Définitions formelles et résultats techniques.....	415
Chapitre XXI.	Estimation de l'erreur d'échantillonnage pour les données d'enquête.....	421
A.	Conceptions d'échantillons d'enquête	422
B.	Questions liées à l'analyse des données d'enquêtes par sondage complexes.....	422
1.	Analyses pondérées.....	422
2.	Aperçu de l'estimation de la variance.....	423
3.	Facteurs de correction d'une population finie (<i>fpc</i>) pour l'échantillonnage sans remplacement	423
4.	Pseudo-strates et pseudo-UPE	424
5.	Approximation courante (<i>WR</i>) pour décrire les plans d'échantillonnage complexes.....	425
6.	Techniques d'estimation de la variance et variables de conception de l'enquête.....	426
7.	Analyse des données d'une enquête par sondage complexe.....	427
C.	Méthodes d'estimation de la variance.....	428
1.	Linéarisation par série de Taylor pour l'estimation de la variance... ..	428
2.	Méthode de réplication pour l'estimation de la variance.....	429
3.	Réplication répétée équilibrée [<i>Balanced repeated replication</i> (BRR)]	430
4.	Techniques de réplication « jackknife » (JK)	430
5.	Quelques erreurs courantes commises par les utilisateurs de logiciels d'estimation de variance	431
D.	Comparaison des logiciels d'estimation de variance.....	432
E.	L'ensemble de données de l'enquête par sondage effectuée au Burundi.....	435
1.	Population étudiée et paramètres de population	435
2.	Plan d'échantillonnage et collecte de données	435
3.	Procédures de pondération et cadre d'estimation de la variance	436
4.	Trois exemples d'analyses de données d'enquêtes	437
F.	Utilisation des procédures d'enquête sans échantillon pour analyser les données d'enquêtes avec échantillonnage.....	437
G.	Les procédures d'enquête par sondage avec SAS 8.2.....	439
1.	Aperçu général de SURVEYMEANS et de SURVEYREG	439
2.	SURVEYMEANS.....	440
3.	SURVEYREG	440

4.	Exemples numériques.....	441
5.	Avantages/inconvénients/coût.....	442
H.	SUDAAN 8.0	442
1.	Aperçu général de SUDAAN	442
2.	DESCRIPT	444
3.	CROSSTAB	445
4.	Exemples numériques.....	445
5.	Avantages/inconvénients/coût.....	446
I.	Procédures d'enquête par sondage dans STATA 7.0	447
1.	Description générale de STATA	447
2.	Svymean, Svyprop, Svytotal, Svylc	448
3.	SVYTAB	449
4.	Exemples numériques.....	449
5.	Avantages/inconvénients/coût.....	450
J.	Procédures d'enquête par sondage dans Epi-Info 6.04d et Epi-Info 2002	450
1.	Aperçu général d'Epi-Info.....	450
2.	Epi-Info Version 6.04d (DOS), module CSAMPLE.....	451
3.	Epi-Info 2002 (Windows).....	452
4.	Exemples numériques.....	453
5.	Avantages/inconvénients/coût.....	453
K.	WesVar 4.2	454
1.	Aperçu général de WesVar	454
2.	Utilisation de WesVar Version 4.2.....	455
3.	Exemples numériques.....	455
4.	Avantages/inconvénients/coût.....	456
L.	PC-CARP	457
M.	CENVAR	458
N.	IVEware (Bêta version)	459
O.	Conclusions et recommandations	459
PARTIE DEUX. Études de cas		465
Introduction		467
Chapitre XXII. Les Enquêtes démographiques et sanitaires		471
A.	Introduction.....	471
B.	Historique	472
C.	Contenu.....	472
D.	Cadre d'échantillonnage	474
E.	Phases de l'échantillonnage	474
F.	Notification des cas de non-réponse.....	476
G.	Comparaison des taux de non-réponse	478
H.	Effets de conception de l'échantillon des DHS	478
I.	Exécution de l'enquête	482
J.	Préparation et traduction des documents de l'enquête	482
K.	Le test préalable.....	483
L.	Recrutement du personnel sur le terrain	484

M.	Formation des enquêteurs	485
N.	Travail de terrain	485
O.	Traitement des données	487
P.	Analyse et rédaction de rapports	489
Q.	Diffusion.....	489
R.	Utilisation des données DHS.....	489
S.	Capacités statistiques.....	490
T.	Enseignements.....	490
Annexe.	Taux de réponse des ménages et des femmes à 66 enquêtes effectuées dans 44 pays, 1990-2000, de différentes régions	494
Chapitre XXIII.	Enquêtes pour l'étude sur la mesure des niveaux de vie.....	497
A.	Introduction.....	497
B.	Pourquoi une enquête LSMS ?	498
C.	Principales caractéristiques des enquêtes LSMS.....	499
1.	Contenu et instruments utilisés.....	499
2.	Questions relatives aux échantillons.....	501
3.	Organisation sur le terrain	502
4.	Qualité	503
5.	Entrée des données.....	505
6.	Durabilité.....	506
D.	Coûts d'exécution d'une enquête LSMS	507
E.	Quel est le degré d'efficacité de la conception des enquêtes LSMS sur leur qualité ?	508
1.	Taux de réponse	508
2.	Non-réponse ponctuelle	509
3.	Vérification de l'homogénéité interne	510
4.	Effets liés à la conception de l'échantillon.....	512
F.	Utilisations des données d'enquêtes LSMS.....	513
G.	Conclusions	514
Annexe I.	Liste des Enquêtes pour l'étude sur la mesure des niveaux de vie.....	516
Annexe II.	Budget des enquêtes LSMS.....	518
Annexe III.	Effet de la conception de l'échantillon sur la précision et l'efficacité des enquêtes LSMS	520
Chapitre XXIV.	Conception de l'enquête et de l'échantillon pour les enquêtes sur le budget des ménages.....	527
A.	Introduction.....	527
B.	Conception de l'enquête	528
1.	Méthodes de collecte de données utilisées dans les enquêtes sur le budget des ménages.....	528
2.	Problèmes de mesure	529
3.	Périodes de référence	530
4.	Fréquence des visites.....	530
5.	Non-réponse	531
C.	Conception de l'échantillon.....	531
1.	Stratification, allocation de l'échantillon aux strates	531
2.	Taille de l'échantillon	532

3.	L'échantillonnage et le choix du moment.....	532
D.	Étude de cas : l'Enquête sur les dépenses et la consommation en République démocratique populaire lao, 1997/98.....	534
1.	Conditions générales d'exécution de l'enquête.....	534
2.	Thèmes traités par l'enquête, les questionnaires	534
3.	Méthodes de mesure	535
4.	Conception de l'échantillon, travail de terrain.....	536
E.	Expériences, leçons apprises.....	536
1.	Méthodes de mesure, non-réponse.....	536
2.	Conception de l'échantillon, erreurs d'échantillonnage	537
3.	Expériences tirées de l'utilisation du journal sur l'emploi du temps.....	538
4.	Utilisation de LECS-2 pour les estimations du PIB.....	538
F.	Conclusion	539
Chapitre XXV.	Les enquêtes sur les ménages dans les pays en transition.....	541
A.	Examen général des enquêtes sur les ménages dans les pays en transition	541
1.	Introduction.....	541
2.	Enquêtes par sondage sur les ménages dans les pays d'Europe centrale et orientale et en URSS avant la période de transition (1991-2000).....	542
3.	Les enquêtes sur les ménages dans la période de transition.....	544
4.	Enquêtes sur le budget des ménages.....	546
5.	Enquêtes sur la main-d'œuvre.....	546
6.	Traits communs de la conception des échantillons et de l'exécution des enquêtes HBS et LFS.....	547
7.	Conclusions.....	555
B.	Les enquêtes sur les ménages dans les pays en transition : études de cas	556
1.	L'enquête par sondage sur les ménages en Estonie.....	556
2.	Conception et exécution de l'Enquête sur le budget des ménages et de l'Enquête sur la main-d'œuvre en Hongrie.....	559
3.	Conception et exécution des enquêtes sur les ménages en Lettonie	563
4.	Les enquêtes par sondage sur les ménages en Lituanie.....	567
5.	Les enquêtes sur les ménages en Pologne pendant la période de transition	569
6.	L'enquête sur la main-d'œuvre et l'enquête sur le budget des ménages en Slovénie.....	575

TABLEAUX

II.1	Effets liés à la conception pour certaines combinaisons de taille d'échantillon en grappes et de corrélation au sein d'un même groupe.....	18
II.2.	Tailles optimales de sous-échantillons pour certaines combinaisons de ratio de coût et de corrélation au sein d'un même groupe	19
II.3.	Erreurs types et intervalles de confiance pour des estimations du taux de pauvreté fondées sur diverses tailles d'échantillon, avec un effet lié à la conception supposé de 2,0	24

II.4.	Coefficient de variation des estimations du taux de pauvreté fondé sur diverses tailles d'échantillon, l'effet supposé lié à la conception étant de 2,0	25
IV.1.	Projet de budget d'une enquête hypothétique sur 3 000 ménages.....	52
VI.1.	Effets liés à la conception d'un échantillonnage disproportionné dans l'hypothèse de deux strates	100
VI.2.	Distributions de la population et trois allocations possibles de l'échantillon à travers les huit provinces (A-H)	113
VII.1.	Caractéristiques des 11 enquêtes sur les ménages incluses dans cette étude.....	120
VII.2.	Estimation des effets liés à la conception découlant de sept enquêtes effectuées en Afrique et en Asie du Sud-Est.....	123
VII. 3.	Estimations des effets liés à la conception pour certaines estimations sur les ménages au niveau du pays et par type de zones géographiques (PNAD 1999).....	124
VII.4.	Estimations des effets liés à la conception pour certaines caractéristiques mesurées au niveau des personnes à l'échelon national et pour divers sous-domaines (PNAD 1999).....	125
VII.5.	Estimation des effets liés à la conception pour certaines estimations de l'enquête PME pour septembre 1999	125
VII.6.	Estimation des effets liés à la conception pour certaines estimations de l'enquête PPV	126
VII.7.	Comparaisons des effets liés à la conception observés dans les différentes enquêtes.....	127
VII.8.	Ensemble des effets liés à la conception divisés en effets résultant de la pondération $d_w^2(\bar{y})$ et de la mise en grappes $d_{cl}^2(\bar{y})$	129
VII.9.	Taux d'homogénéité pour les domaines urbain et rural.....	130
X.1.	Liste sommaire des déterminants de la qualité de l'échantillonnage	198
X.2.	Liste récapitulative des procédures d'examen de la traduction	200
X.3.	Liste récapitulative à utiliser pour l'examen des procédures de formation	203
X.4.	Liste récapitulative des actions entrant dans l'exécution d'une enquête	206
X.5.	Liste récapitulative des vérifications à faire en ce qui concerne le processus d'entrée des données	209
XI.1.	Quelques caractéristiques des principales enquêtes par sondage sur les ménages au Brésil.....	223
XI.2.	Estimation des taux d'omission dans les recensements de la population du Brésil découlant des enquêtes après les recensements de 1991 et 2000 (<i>en pourcentage</i>)	226
XIII.1.	Durée estimative du travail de terrain dans un village	261
XIII.2.	Estimation des coûts de LECS-3 (<i>en dollars des États-Unis par jour</i>)	261
XIII.3.	Tailles optimales de l'échantillon (<i>mopt</i>) et efficacité relative de la conception ($m = 15$) pour différentes valeurs de ρ	263
XIV.1.	Proposition de calendrier pour une enquête sur le secteur informel.....	267
XIV. 2.	Matrice des postes comptables en regard des activités d'enquête	270

XIV.3.	Matrice du temps de personnel (jours) en regard des activités de l'enquête	271
XIV.4.	Coûts par poste comptable en proportion du budget total : enquêtes sur les objectifs de la fin de la décennie (1999-2000), pour certains pays africains (<i>en pourcentage</i>)	274
XIV.5.	Proportion du budget allouée aux différents postes comptables : évaluation de l'impact des services aux microentreprises [Assessing the Impact of Macroenterprise Services (AIMS), Zimbabwe (1999)] (<i>en pourcentage</i>)	275
XIV.6.	Coûts des différentes activités d'une enquête en proportion de son budget total : objectifs de la fin de la décennie (1999-2000), pour certains pays africains (<i>en pourcentage</i>)	277
XIV.7.	Coût des activités de l'enquête en proportion de son budget total : AIMS, Zimbabwe (1999) [<i>en pourcentage</i>]	277
XIV.8.	Coûts par poste comptable pour chaque activité de l'enquête en proportion prévue du budget : AIMS, Zimbabwe (1999) [<i>en pourcentage</i>]	278
XIV.9.	Coûts par poste comptable pour chaque activité de l'enquête en proportion effective du budget : AIMS, Zimbabwe (1999) [<i>en pourcentage</i>]	279
XV.1.	Données d'une enquête sur les ménages mises en mémoire dans un simple fichier rectangulaire	302
XVI.1.	Distribution de la population par âge et par sexe, Saipan, Commonwealth des îles Mariannes du Nord, avril 2002 (<i>pourcentages par rangée</i>)	322
XVI.2.	Distribution de la population par âge et par sexe, Saipan, Commonwealth des îles Mariannes du Nord, avril 2002 (<i>pourcentages par colonne</i>)	322
XVI.3.	Statistiques sommaires sur le revenu des ménages par groupe ethnique, Samoa américaine, 1994	323
XVI.4.	Types d'éclairage utilisés par les familles vietnamiennes, 1991-1993	324
XVI.5.	Données récapitulatives sur les dépenses totales des ménages : Viet Nam, 1992-1993 (<i>en milliers de dongs par an</i>)	327
XVI.6.	Utilisation des services de santé par les personnes (de tous âges) qui se sont rendues dans un établissement de soins au cours des quatre dernières semaines, par zones urbaines et rurales du Viet Nam, en 1992-1993	328
XVI.7.	Dépenses totales des ménages par région au Viet Nam, 1992-1993 (<i>en milliers de dong par an</i>)	329
XVIII.1.	Quelques techniques multivariates et leur but	354
XVIII.2.	Données relatives aux exploitations indiquant la présence ou l'absence d'un ensemble de caractéristiques communes	359
XVIII.3.	Matrice de similarité entre exploitations	359
XVIII.4.	Résultats d'une analyse en composantes principales	361
XVIII.5.	Variables utilisées (et leur pondérations) pour la construction d'un indice de prévision des dépenses de consommation dans la région du Kilimandjaro, en République-Unie de Tanzanie	365
XVIII.6.	Points de coupure utilisés pour séparer la population en cinq quintiles de richesse	367
XIX.1.	Structure type de la conception d'une enquête sur les ménages	372

XIX.2.	Interprétation des estimations des paramètres de régression linéaire lorsque la variable dépendante correspond aux revenus salariaux du ménage pour le modèle 1.....	383
XIX.3.	Revenus salariaux des ménages qui peuvent être estimés (modèle 1)....	383
XIX.4.	Interprétation des estimations des paramètres de régression linéaire lorsque la variable dépendante indique les revenus salariaux des ménages, selon le modèle 2.....	384
XIX.5.	Interprétation des estimations des paramètres de régression logistique lorsque la variable dépendante est un indicateur des ménages se situant en dessous du seuil de pauvreté, avec le modèle 4.....	387
XX.1.	Distorsion et carré moyen de l'estimateur des moindres carrés ordinaires et variances des estimateurs non biaisés pour une population de 3 850 exploitations, déterminés à partir de diverses conceptions d'enquêtes.....	406
XX.2.	Tableau ANOVA comparant les régressions pondérées et non pondérées	407
XX.3.	Ratios de trois tests de chi-carré itérés à des tests de SRS	410
XX.4.	Estimation des tailles asymptotiques de tests fondés sur X^2 et sur X_C^2 de certains éléments de l'Enquête générale de 1971 sur les ménages du Royaume-Uni de Grande-Bretagne et d'Irlande du Nord; la taille normale est de 0,05.....	410
XX.5.	Dimensions asymptotiques estimatives de tests fondés sur $X_I^2, X_I^2/\delta^2$, et sur $X_I^2/\hat{\lambda}^2$ pour une classification croisée de certaines variables tirées de l'Enquête générale de 1971 sur les ménages du Royaume-Uni de Grande-Bretagne et d'Irlande du Nord; la taille normale est de 0,05	411
XX.6.	Niveaux estimatifs de signification asymptotique (SL) de X^2 et des statistiques corrigées $X^2/\hat{\delta}^2$, $X^2/\hat{\lambda}^2$, X^2/\hat{d}^2 : tableau 2 x 5 x 4 et niveau de signification nominale $\alpha = 0,05$	413
XXI.1.	Comparaison de PROCS selon cinq logiciels : pourcentage estimatif et nombre de femmes qui sont séropositives, et estimation de l'erreur type, femmes ayant accouché récemment, au Burundi, 1988-1989	432
XXI.2.	Attributs de huit logiciels permettant l'estimation de la variance pour les données d'enquête par sondage complexes.....	434
XXII.1.	Valeurs moyennes $d(\bar{y})$ et $\hat{\rho}$ pour 48 enquêtes DHS, 1984-1993	481
XXIII.1.	Contenu du questionnaire de l'enquête sur les ménages au Viet Nam, 1997-1998	500
XXIII.2.	Exemples de modules additionnels.....	501
XXIII.3.	Contrôles de qualité dans les enquêtes LSMS	504
XXIII.4.	Taux de réponse enregistrés lors de récentes enquêtes LSMS	509
XXIII.5.	Fréquence des données manquantes dans les enquêtes LSMS et les enquêtes LFS	510
XXIII.6.	Ménages avec agrégat complet de consommation : exemples tirés de récentes enquêtes LSMS	511
XXIII.7.	Homogénéité interne des données : liens réussis entre les modules (en pourcentage).....	511
XXIII.8.	Exemples d'effets de conception dans les enquêtes LSMS surveys.....	512
AIII.1.	Variation des effets de conception par variable, Ghana, 1987.....	521
AIII.2.	Variation des effets de conception dans le temps, Ghana, 1987 et 1988.....	522

AIII.3.	Variation des effets de conception entre pays	523
AIII.4.	Description des variables de l'analyse : niveau individuel.....	524
AIII.5.	Description des variables de l'analyse : niveau des ménages.....	534
XXIV.1.	Effets de conception sur la consommation des ménages et sur la possession de biens durables.....	537
XXIV.2.	Ratio entre le nombre effectif et le nombre attendu de personnes dans l'échantillon pour le journal sur l'emploi du temps.....	538
XXV.1.	Nouvelles enquêtes sur le budget des ménages et sur la main-d'œuvre dans les pays en transition, 1992-2000 : première année, périodicité et année de la dernière refonte	545
XXV.2.	Taille de l'échantillon et conception et méthodes d'estimation de l'échantillon dans les enquêtes HBS et LFS de 2000 de certains pays en transition.....	551
XXV.3.	Taux de non-réponse à l'enquête HBS dans certains pays en transition, 1992-2000.....	552
XXV.4.	Taux de non-réponse à l'enquête LFS dans certains pays en transition, 1992-2000	553
XXV.5.	Structure des coûts de l'enquête HBS de 2000 en Hongrie	554
XXV.6.	Structure des coûts de l'enquête LFS de 2000 en Hongrie.....	555

FIGURES

III.1.	Illustration du formatage d'un questionnaire	41
IV.1.	Plan de travail pour l'élaboration et l'exécution d'une enquête sur les ménages.....	54
X.1.	Procédures d'assurance de qualité de l'ESM	192
X.2.	Parcours des données de l'ESM et mesures d'assurance de la qualité correspondantes.....	207
X.3.	Exemple d'indice de déviation d'un échantillon.....	212
XV.1.	Enquête sur les niveaux de vie au Népal II.....	304
XV.2.	Utilisation d'un tableur comme cadre d'échantillonnage de la première phase.....	306
XV.3.	Exécution d'une stratification implicite	307
XV.4.	Sélection d'un échantillon PPS (première étape)	308
XV.5.	Sélection d'un échantillon PPS (deuxième étape).....	309
XV.6.	Sélection d'un échantillon PPS (troisième étape).....	310
XV.7.	Sélection d'un échantillon PPSA (quatrième étape)	311
XV.8.	Tableur indiquant les unités primaires d'échantillonnage sélectionnées	312
XV.9.	Calcul des probabilités de la sélection de la première phase	313
XV.10.	Résultats de l'établissement des listes de ménages	314
XV.11.	Non-réponse.....	315
XV.12.	Calcul des probabilités et des pondérations de la deuxième phase	315
XVI.1.	Types d'éclairage utilisés par les familles vietnamiennes, 1992-1993 (graphique).....	325
XVI.2.	Types d'éclairage utilisés par les familles vietnamiennes, 1992-1993 (diagramme circulaire) [<i>en pourcentage</i>].....	325
XVI.3.	Distribution par âge de la population de Saipan, avril 2002 (histogramme).....	326

XVI.4.	Utilisation des services de santé par les personnes (de tous âges) qui se sont rendues dans un établissement de soins au cours des quatre dernières semaines, par zones urbaines et rurales du Viet Nam, en 1992-1993 (<i>en pourcentage</i>).....	328
XVIII.1.	Exemple de représentation matricielle de 4 variables.....	356
XVIII.2.	Dendogramme formé d'après la matrice de similarité entre exploitations.....	360
XIX.1.	Application de pondérations et estimation statistique.....	374
XX.1.	Pas de sélection.....	399
XX.2.	Sélection sur X : $XL < X < XU$	399
XX.3.	Sélection sur X : $X < XL$; $X > XU$	399
XX.4.	Sélection sur Y : $YL < Y < YU$	399
XX.5.	Sélection sur Y : $Y < YL$; $Y > YU$	399
XX.6.	Sélection sur Y : $Y > YU$	399
XXIII.1.	Rapport entre l'objet des enquêtes LSMS et les instruments de l'enquête.....	499
XXIII.2.	Calendrier mensuel d'activités de chaque équipe.....	503
XXIII.3.	Éléments de coût d'une enquête LSMS (part du coût total).....	507

Liste des experts ayant contribué à la réalisation de cette publication

Participants à la réunion du groupe d'experts sur les caractéristiques opérationnelles des enquêtes sur les ménages dans les pays en développement et les pays en transition (8-10 octobre 2002, New York)

Savitri Abeyasekera
Université de Reading
Reading, Royaume-Uni de Grande-
Bretagne et d'Irlande du Nord

Oladejo O. Ajayi
Consultant en statistiques
Ikoyi, Lagos, Nigéria

Jeremiah Banda
DESA/UNSD
New York, New York
États-Unis d'Amérique

Grace Bediako
DESA/UNSD
New York, New York
États-Unis d'Amérique

Donna Brogan
University d'Emory
Atlanta, Georgia
États-Unis d'Amérique

Mary Chamie
DESA/UNSD
New York, New York
États-Unis d'Amérique

James R. Chromy
Research Triangle Institute
Research Triangle Park
Caroline du Nord,
États-Unis d'Amérique

Willem de Vries
DESA/UNSD
New York, New York
États-Unis d'Amérique

Paul Glewwe
Université du Minnesota
St. Paul, Minnesota
États-Unis d'Amérique

Ivo Havinga
DESA/UNSD
New York, New York
États-Unis d'Amérique

Rosaline Hirschowitz
Statistics South Africa
Pretoria, Afrique du Sud

Gareth Jones
Fonds des Nations Unies pour l'enfance
New York, New York
États-Unis d'Amérique

Graham Kalton
Westat
Rockville, Maryland
États-Unis d'Amérique

Hiroshi Kawamura
DESA/Division de l'analyse des
politiques de développement
Nations Unies
New York, New York
États-Unis d'Amérique

Erica Keogh
Université du Zimbabwe
Harare, Zimbabwe

Jan Kordos
École d'économie de Varsovie
Varsovie, Pologne

James Lepkowski
Institute for Social Research
Ann Arbor, Michigan
États-Unis d'Amérique

Gad Nathan
Hebrew University
Jérusalem, Israël

Frederico Neto
 DESA/Division de l'analyse
 des politiques de développement
 Nations Unies
 New York, New York
 États-Unis d'Amérique

Colm O'Muircheartaigh
 University of Chicago
 Chicago, Illinois
 États-Unis d'Amérique

Hans Pettersson
 Statistics Sweden
 Stockholm, Suède

Hussein Sayed
 Université du Caire
 Orman, Giza, Égypte

Michelle Schoch
 Fonds des Nations Unies
 pour la population
 New York, New York
 États-Unis d'Amérique

Stefan Schweinfest
 DESA/UNSD
 New York, New York
 États-Unis d'Amérique

Anatoly Smyshlyev
 DESA/Division de l'analyse
 des politiques de développement
 Nations Unies
 New York, New York
 États-Unis d'Amérique

Pedro Silva
 Função Instituto Brasileiro
 de Geografia e Estatística
 Rio de Janeiro, Brésil

Diane Steele
 Banque mondiale
 Washington, D.C.
 États-Unis d'Amérique

Sirageldin Suliman
 DESA/UNSD
 New York, New York
 États-Unis d'Amérique

T. Bedirhan Üstün
 Organisation mondiale de la santé
 Genève, Suisse

Shyam Upadhyaya
 Integrated Statistical Services
 (INSTAT)
 Kathmandu, Népal

Martin Vaessen
 Programme d'enquêtes démographiques
 et sanitaires
 ORC Macro*
 Calverton, Maryland
 États-Unis d'Amérique

Ibrahim Yansaneh
 Commission de la fonction publique
 internationale
 [DESA/UNSD]
 New York, New York
 États-Unis d'Amérique

* Société de conseil en sondages
 d'opinion.

Auteurs

Savitri Abeyasekera
University de Reading
Reading,
Royaume-Uni de Grande-Bretagne
et d'Irlande du Nord

J. Michael Brick
Westat
Rockville, Maryland
États-Unis d'Amérique

Donna Brogan
Université d'Emory
Atlanta, Georgia
États-Unis d'Amérique

Somnath Chatterji
Organisation mondiale de la santé
Genève, Suisse

James R. Chromy
Research Triangle Institute
Research Triangle Park
Caroline du Nord,
États-Unis d'Amérique

Paul Glewwe
Université du Minnesota
St. Paul, Minnesota
États-Unis d'Amérique

Hermann Habermann
United States Census Bureau
Suitland, Maryland
États-Unis d'Amérique

Graham Kalton
Westat
Rockville, Maryland
États-Unis d'Amérique

Daniel Kasprzyk
Mathematica Policy Research
Washington, D.C.,
États-Unis d'Amérique

Erica Keogh
Université du Zimbabwe
Harare, Zimbabwe

Jan Kordos
École d'économie de Varsovie
Varsovie, Pologne

Thanh Lê
Westat
Rockville, Maryland
États-Unis d'Amérique

James Lepkowski
Université du Michigan
Ann Arbor, Michigan
États-Unis d'Amérique

Michael Levin
United States Census Bureau
Washington, D.C.
États-Unis d'Amérique

Abdelhay Mechbal
Organisation mondiale de la santé
Genève, Suisse

Juan Muñoz
Consultant indépendant
Santiago, Chili

Christopher J. L. Murray
Organisation mondiale de la santé
Genève, Suisse

Gad Nathan
Hebrew University
Jérusalem, Israël

Hans Pettersson
Statistics Sweden
Stockholm, Suède

Kinnon Scott
Banque mondiale
Washington, D.C.
États-Unis d'Amérique

Pedro Silva
Função Instituto Brasileiro
de Geografia e Estatística (IBGE)
Rio de Janeiro, Brésil

Bounthavy Sisouphantong
National Statistics Centre
Vientiane, République populaire
démocratique lao

Diane Steele
Banque mondiale
Washington, D.C.
États-Unis d'Amérique

Tilahun Temesgen
Banque mondiale
Washington, D.C.
États-Unis d'Amérique

Mamadou Thiam
Organisation des Nations Unies pour
l'éducation, la science et la culture
Montréal, Canada

T. Bedirhan Üstun
Organisation mondiale de la santé
Genève, Suisse

Martin Vaessen
Programme d'enquêtes
démographiques et sanitaires
ORC Macro*

Calverton, Maryland
États-Unis d'Amérique

Vijay Verma
Université de Sienne
Sienna, Italie

Ibrahim Yansaneh
Commission de la fonction publique
internationale
[DESA/UNSD]
New York, New York
États-Unis d'Amérique

* Société de conseil en sondages
d'opinion.

Éditeurs

Oladejo Ajayi
Consultant en statistiques
Lagos, Nigéria

Paul Biemer
Research Triangle Institute
Research Triangle Park
Caroline du Nord,
États-Unis d'Amérique

Steven B. Cohen
Agency for Healthcare Research
and Quality
Rockville, Maryland
États-Unis d'Amérique

John Eltinge
United States Bureau of Labor Statistics
Washington, D.C.
États-Unis d'Amérique

Paul Glewwe
Université du Minnesota
St. Paul, Minnesota
États-Unis d'Amérique

Barry Graubard
National Cancer Institute
Bethesda, Maryland
États-Unis d'Amérique

Stephen Haslett
Université de Massey
Palmerston North
Nouvelle-Zélande

Steven Heeringa
Université du Michigan
Ann Arbor, Michigan
États-Unis d'Amérique

Thomas B. Jabine
Consultant en statistiques
Washington, D.C.
États-Unis d'Amérique

Gareth Jones
Fonds des Nations Unies pour l'enfance
New York, New York

William D. Kalsbeek
Université de Caroline du Nord

Chapel Hill, Caroline du Nord
États-Unis d'Amérique

Graham Kalton
Westat
Rockville, Maryland
États-Unis d'Amérique

Ben Kiregyera
Uganda Bureau of Statistics
Kampala, Ouganda

Jan Kordos
École d'économie de Varsovie
Varsovie, Pologne

Phil Kott
United States Department of Agriculture
National Agricultural Statistics Service
Fairfax, Virginie
États-Unis d'Amérique

Karol Krotki
NuStats
Austin, Texas
États-Unis d'Amérique

James Lepkowski
Université du Michigan
Ann Arbor, Michigan
États-Unis d'Amérique

Dalisay Maligalig
Banque asiatique de développement
Manille, Philippines

David Marker
Westat
Rockville, Maryland
États-Unis d'Amérique

Juan Muñoz
Consultant indépendant
Santiago, Chili

Gad Nathan
Hebrew University
Jérusalem, Israël

Colm O'Muircheartaigh
Université de Chicago
Chicago, Illinois
États-Unis d'Amérique

Robert Pember
Organisation internationale du Travail
Bureau de statistique
Genève, Suisse

Robert Santos
NuStats
Austin, Texas
États-Unis d'Amérique

Pedro Silva
Função Instituto Brasileiro de
Geografia e Estatística (IBGE)
Rio de Janeiro, Brésil

Anthony G. Turner
Consultant en échantillonnage
Jersey City, New Jersey
États-Unis d'Amérique

Ibrahim Yansaneh
Commission de la fonction publique
internationale
[DESA/UNSD]
New York, New York
États-Unis d'Amérique

Première partie

**Conception,
mise en œuvre
et analyse des enquêtes**

Chapitre premier

Introduction

IBRAHIM S. YANSANEH*

Commission de la fonction publique internationale
Nations Unies, New York

* Ancien chef, Service de la méthodologie et de l'analyse, DESA/UNSD.

RÉSUMÉ

Le présent chapitre offre un bref aperçu des enquêtes sur les ménages effectuées dans les pays en développement et les pays en transition. Il indique également les objectifs généraux de cette publication et l'importance pratique de ces objectifs.

Termes clés : enquêtes sur les ménages, caractéristiques opérationnelles, conception d'une enquête complexe, coûts des enquêtes, erreurs commises dans le cadre des enquêtes.

A. Enquêtes sur les ménages dans les pays en développement et les pays en transition

1. Au cours des dernières décennies, on a assisté à un accroissement de la demande de données démographiques et socio-économiques sur les ménages et les individus dans les pays en développement et les pays en transition. Ces données sont devenues indispensables pour l'analyse des politiques économiques et sociales, la planification du développement, la gestion des programmes et la prise de décisions à tous les niveaux. Pour répondre à cette demande, les décideurs et autres parties prenantes ont souvent recours aux enquêtes sur les ménages. C'est pourquoi ces enquêtes sont devenues l'un des principaux mécanismes de collecte d'informations sur les populations des pays en développement et des pays en transition. Elles jouent à présent un rôle central et stratégique dans l'organisation des systèmes nationaux de statistiques et la formulation de politiques. Aujourd'hui, la plupart des pays ont des systèmes de collecte de données qu'ils utilisent pour leurs enquêtes sur les ménages, mais ils ne présentent pas tous les mêmes niveaux d'expérience et d'infrastructure. Les enquêtes effectuées par les offices nationaux de statistique sont généralement polyvalentes ou intégrées, et conçues pour fournir des données fiables sur un ensemble de caractéristiques démographiques et socio-économiques de leurs diverses populations. Les enquêtes sur les ménages sont également utilisées pour étudier les petites et moyennes entreprises et les petites exploitations agricoles des pays en développement et des pays en transition.

2. Outre les enquêtes nationales financées sur le budget ordinaire des pays, un grand nombre d'enquêtes sur les ménages sont effectuées dans les pays en développement et les pays en transition sur l'initiative d'institutions internationales, afin d'établir et de suivre des estimations nationales des caractéristiques ou des indicateurs qui intéressent ces institutions, et de comparer ces indicateurs à l'échelle internationale. La plupart de ces enquêtes sont effectuées de façon ad hoc, mais on note un regain d'intérêt pour l'établissement de programmes permanents d'enquêtes intégrées à thèmes multiples et à cycles multiples de la part d'organisations internationales telles que l'Or-

ganisation des Nations Unies et la Banque mondiale, à tous les stades de la conception, de la mise en œuvre, de l'analyse de ces enquêtes et de la diffusion de leurs résultats. Les principaux exemples d'enquêtes sur les ménages effectuées par des institutions internationales dans les pays en développement sont les Enquêtes démographiques et sanitaires (*Demographic and Health Surveys*, DHS) effectuées par ORC Macro pour l'United States Agency for International Development (USAID); le programme d'étude sur la mesure des niveaux de vie (*Living Standards Measurement Study*, LSMS), exécuté avec l'assistance technique de la Banque mondiale; et les Enquêtes en grappes à indicateurs multiples (MICS) effectuées par le Fonds des Nations Unies pour l'enfance (UNICEF). Ces programmes d'enquêtes se déroulent dans divers pays en développement d'Afrique, d'Asie, d'Amérique latine et des Caraïbes et du Moyen-Orient. Les programmes d'enquêtes DHS et LSMS sont décrits de façon détaillée dans les études de cas présentées respectivement aux chapitres V et VI. Voir également Banque mondiale (2000) pour une analyse détaillée d'autres programmes d'enquêtes exécutés par la Banque mondiale dans les pays en développement, y compris les Enquêtes prioritaires et les enquêtes effectuées à l'aide du Questionnaire sur les principaux indicateurs de bien-être (*Core Welfare Indicators Questionnaire*, CWIQ). Pour plus de détails sur les MICS, voir UNICEF (2000). Le programme DHS est l'héritier d'un ancien programme d'enquêtes, l'Enquête mondiale sur la fécondité (*World Fertility Survey*, WFS), financée conjointement par l'USAID et le Fonds des Nations Unies pour la population (FNUAP), avec l'aide des gouvernements du Royaume-Uni de Grande-Bretagne et d'Irlande du Nord, des Pays-Bas et du Japon. Pour plus de détails sur le programme WFS, voir Verma *et al.* (1980).

B. Objectifs de la présente publication

3. La présente publication offre un cadre méthodologique pour l'exécution d'enquêtes dans les pays en développement et les pays en transition. Face au grand nombre d'enquêtes effectuées dans ces pays, il y a un besoin constant de travail méthodologique à tous les stades du processus d'enquête et d'utilisation des meilleures méthodes par les producteurs et les utilisateurs des données des enquêtes sur les ménages. Une large part de ce travail méthodologique s'effectue sous les auspices d'institutions internationales et de DESA/UNSD, à travers ses publications et ses rapports techniques. Cette publication est le tout dernier fruit de ces travaux.

4. La plupart des enquêtes effectuées aujourd'hui dans les pays en développement et les pays en transition se fondent sur la méthodologie et les procédures courantes appliquées dans le monde entier. Toutefois, un grand nombre de ces enquêtes s'inscrivent dans le contexte des graves difficultés budgétaires auxquelles se heurtent des pays qui offrent des niveaux très divers d'infrastructure et de capacité technique d'exécution. On constate un très net besoin non seulement de développement et de progrès constants des méthodologies employées, mais aussi de transfert de ces méthodologies aux pays en développement et aux pays en transition. Le meilleur moyen de répondre à ce besoin est de recourir à la coopération technique et au renforcement des capacités statistiques. Cette publication, qui a pour objet de servir d'instrument à ce renforcement des capacités, constitue une source centrale de données techniques et autres informations nécessaires à une conception et à une mise en œuvre efficaces des enquêtes sur les ménages et à la bonne utilisation des données recueillies.

5. Cette publication s'adresse à toutes les personnes engagées dans la production et l'utilisation de données d'enquêtes, y compris :

- Les membres du personnel des offices nationaux de statistiques;

- Les consultants internationaux qui dispensent une assistance technique aux pays;
- Les chercheurs et autres analystes qui analysent les données des enquêtes sur les ménages;
- Les personnes qui enseignent et celles qui étudient les méthodes d'enquête.

6. Cette publication offre une source complète de données et de références sur les aspects importants de la conception, de la mise en œuvre et de l'analyse des enquêtes par échantillonnage sur les ménages dans les pays en développement et les pays en transition. Le lecteur peut utiliser les informations méthodologiques générales présentées dans la première partie de cet ouvrage, avec les études de cas de la deuxième partie, pour concevoir de nouvelles enquêtes à effectuer dans ces pays. En particulier, cette publication a pour buts :

- a) D'offrir une source centrale de données et de références sur les aspects techniques de la conception, la mise en œuvre et l'analyse des enquêtes dans les pays en développement et les pays en transition;
- b) D'aider les praticiens à concevoir et mettre en œuvre plus efficacement les enquêtes sur les ménages;
- c) De présenter des études de cas sur divers types d'enquêtes qui ont été ou sont effectuées dans certains pays en développement ou en transition, en mettant l'accent sur les éléments qui peuvent être généralisés, afin d'aider les praticiens à concevoir et mettre en œuvre de nouvelles enquêtes dans les mêmes pays ou dans d'autres;
- d) D'examiner plus en détail certains éléments de trois caractéristiques opérationnelles des enquêtes — les effets liés à la conception, les coûts et les erreurs autres que les erreurs d'échantillonnage — et d'évaluer les possibilités d'application de ces caractéristiques ou de leurs éléments à différentes enquêtes et différents pays;
- e) D'offrir des directives pratiques pour l'analyse des données provenant d'enquêtes par échantillonnage complexes et une comparaison détaillée des types de logiciels utilisables pour cette analyse.

C. Importance pratique de ces objectifs

7. Les enquêtes sur les ménages effectuées dans les pays en développement et les pays en transition présentent de nombreux traits communs. De plus, il y a souvent des similarités entre pays, surtout entre pays d'une même région, en ce qui concerne les principales caractéristiques des populations étudiées. Dans la mesure où les échantillons retenus pour les enquêtes sur les ménages et les caractéristiques des populations étudiées sont similaires d'un pays à un autre, nous pouvons nous attendre à ce que certaines caractéristiques opérationnelles ou certains de leurs éléments soient également similaires, ou transférables d'un pays à un autre.

8. La possibilité de transfert des caractéristiques opérationnelles des enquêtes offre plusieurs avantages pratiques. Premièrement, les informations sur la conception d'une enquête donnée dans un pays particulier peuvent contenir des conseils pratiques propres à améliorer l'efficacité de la même enquête lorsqu'elle se répète dans le même pays ou d'une enquête similaire effectuée dans un autre pays. Deuxièmement, les pays disposant d'une infrastructure insuffisante ou dépassée pour effectuer une enquête peuvent tirer un immense profit de données empiriques sur les caractéristiques de conception et de mise en œuvre d'un échantillon provenant d'autres pays dotés d'une meilleure infrastructure d'enquête et d'une meilleure capacité statistique générale. Troisièmement, le fait que des informations liées à la conception d'un échantillon

coûteux peuvent être « empruntées » à une enquête précédente offre de substantielles possibilités d'économies. De plus, l'expérience pratique découlant d'une enquête précédente peut contribuer à maximiser l'efficacité de conception de l'enquête envisagée.

9. Outre le fait qu'elle traite des questions de coût et d'efficacité de la conception et de la mise en œuvre d'enquêtes, cette publication a pour objectif général non négligeable de favoriser l'élaboration d'enquêtes de qualité sur les ménages dans les pays en développement et les pays en transition. Elle s'inspire de précédentes initiatives des Nations Unies, telles que le Programme de mise en place de dispositifs nationaux d'enquête sur les ménages (NHSCP), qui s'est achevé il y a plus de dix ans. Les études de cas offrent des directives importantes sur les aspects de la conception et de la mise en œuvre d'enquêtes qui ont donné de bons résultats dans les pays en développement et les pays en transition, sur les pièges à éviter et sur les mesures à prendre pour améliorer la fiabilité des données recueillies et réduire les coûts totaux des enquêtes. Le fait que toutes les enquêtes décrites ici ont été effectuées dans des pays en développement ou des pays en transition rend cette publication d'autant plus pertinente et utile pour le développement des statistiques dans ces pays.

10. L'analyse et la diffusion des données d'enquêtes sont parmi les domaines qui ont le plus besoin de renforcement des capacités dans les pays en développement et les pays en transition. Les analyses des données des enquêtes vont rarement au-delà des fréquences et des tabulations de base. Grâce à une analyse appropriée des données d'enquêtes et à la diffusion en temps opportun des résultats de cette analyse, les informations requises sont immédiatement accessibles pour la formulation de politiques et les décisions en matière d'affectation des ressources. Cette publication fournit des indications pratiques sur la manière de procéder à des analyses plus élaborées de microdonnées, sur les moyens de tenir compte de la complexité de la conception dans l'analyse des données engendrées, sur la façon d'incorporer les objectifs de l'analyse au stade de la conception et sur les moyens d'utiliser des logiciels spéciaux pour analyser les données d'enquêtes complexes.

11. En résumé, cette publication constitue une source exhaustive de références sur tous les aspects des enquêtes sur les ménages effectuées dans les pays en développement et les pays en transition. Les descriptions techniques contenues dans la première partie, combinées aux exemples concrets et aux études de cas de la deuxième partie, devraient se révéler utiles aux praticiens des enquêtes à travers le monde pour la conception, la mise en œuvre et l'analyse des nouvelles enquêtes sur les ménages.

RÉFÉRENCES

- Banque mondiale (2000). Poverty in Africa: survey databank. Accessible sur <http://www4.worldbank.org/afr/poverty>.
- Fonds des Nations Unies pour l'enfance (UNICEF) [2000]. Manuel établi en fin de décennie sur la conduite des enquêtes en grappes à indicateurs multiples. New York, UNICEF, février.
- Verma, V., C. Scott et C. O'Muircheartaigh (1980). Modèles d'échantillons et erreurs d'échantillonnage pour l'Enquête mondiale sur la fécondité. *Journal of the Royal Statistical Society, Séries A*, vol. 143, pp. 431-473. Avec discussion.

Section A

**Conception
et mise en œuvre
des enquêtes**

Chapitre II

Aperçu des problèmes de conception d'échantillons pour les enquêtes sur les ménages dans les pays en développement et les pays en transition

IBRAHIM S. YANSANEH*

Commission de la fonction publique internationale
Nations Unies, New York

* Ancien chef, Service de la méthodologie et de l'analyse, DESA/UNSD.

RÉSUMÉ

Le présent chapitre traite des problèmes clés que pose la conception d'échantillons nationaux, principalement pour les enquêtes sur les ménages dans les pays en développement et les pays en transition. Il couvre certains thèmes tels que les cadres d'échantillonnage, la taille des échantillons, l'échantillonnage stratifié à phases multiples, l'estimation du domaine et l'analyse de l'enquête. En outre, ce chapitre sert d'introduction à toutes les phases du processus d'enquête qui sont traitées de façon détaillée tout au long de cette publication, tout en soulignant le lien entre chacune de ces phases et le processus de conception de l'échantillon.

Termes clés : conception d'un échantillon complexe, cadre d'échantillonnage, population cible, stratification, échantillonnage en grappes, unité primaire d'échantillonnage.

A. Introduction

1. Modèles d'échantillon pour les enquêtes dans les pays en développement et les pays en transition

1. Le présent chapitre offre un aperçu des problèmes liés à la conception d'échantillons nationaux pour les enquêtes sur les ménages dans les pays en développement et les pays en transition. Comme l'ensemble de cette publication, il se concentre sur les enquêtes sur les ménages. Les enquêtes sur les entreprises et l'agriculture ne sont pas traitées explicitement, mais une grande partie des informations données leur sont également applicables.

2. Les modèles d'échantillons pour les enquêtes sur les ménages dans les pays en développement et les pays en transition présentent beaucoup de traits communs. La plupart des enquêtes utilisent des modèles de probabilité fondés sur un échantillonnage stratifié à phases multiples. Ces modèles sont utilisés principalement pour l'élaboration de cadres et pour le regroupement des entrevues en grappes afin de réduire

leur coût. Le choix de l'échantillon se fait généralement à l'intérieur d'une même strate (voir section B). Les unités sélectionnées lors de la première phase, dénommées unités primaires d'échantillonnage (UPE) [en anglais, *primary sampling units*, UPE) dans les écrits sur les enquêtes par échantillonnage, sont souvent constituées à partir des zones d'énumération identifiées et utilisées lors d'un précédent recensement national de la population et du logement. Il peut s'agir de quartiers urbains ou de villages des zones rurales. Dans certains pays, les UPE peuvent être les quartiers des responsables du recensement, des districts administratifs ou leurs subdivisions. Les unités choisies à l'intérieur de chaque UPE sélectionnée sont appelées unités secondaires; les unités choisies au troisième stade sont appelées unités tertiaires, et ainsi de suite. Pour les enquêtes sur les ménages des pays en développement et des pays en transition, les unités secondaires sont généralement des logements ou des ménages, et les unités sélectionnées au troisième stade sont généralement des personnes. En général, les unités choisies au dernier stade d'un échantillonnage à phases multiples sont appelées les unités d'échantillonnage ultimes.

3. Malgré les nombreuses similarités évoquées ci-dessus, les modèles d'échantillon utilisés pour les enquêtes dans les pays en développement et les pays en transition ne sont pas identiques d'un pays à l'autre, et peuvent varier, par exemple, du point de vue de la population cible, du contenu et des objectifs, du nombre de strates, des taux d'échantillonnage à l'intérieur d'une même strate, de la taille des échantillons au sein des UPE et du nombre d'UPE choisies au sein d'une même strate. En outre, la population concernée peut varier du point de vue des taux de présence de certaines caractéristiques démographiques, du degré d'hétérogénéité au sein d'une même strate et d'une strate à une autre, et de la répartition de certaines sous-catégories de population au sein d'une même strate et d'une strate à une autre.

2. Aperçu général

4. Ce chapitre est articulé de la façon suivante : la section A est une introduction générale. La section B est consacrée à un examen des modèles d'échantillon stratifiés à phases multiples. Tout d'abord, elle décrit l'échantillonnage à probabilité proportionnelle à la taille de l'échantillon. Puis elle présente la notion d'effet lié à la conception dans le cadre de l'échantillonnage en grappes. Vient ensuite une discussion sur les choix optimaux concernant le nombre d'UPE et le nombre d'unités secondaires (logements, ménages, personnes, etc.) à l'intérieur des UPE. Cette discussion tient compte de certains facteurs tels que la précision exigée des estimations des enquêtes et certaines considérations pratiques découlant de l'organisation du travail de terrain. La section C traite des cadres d'échantillonnage et des problèmes connexes. La section D porte sur la question de l'estimation par domaine et des diverses formules d'allocation qui peuvent être envisagées pour répondre aux demandes concurrentes découlant du désir d'établir des estimations aux niveaux national et sous-national. La section E examine la question de la détermination de la taille de l'échantillon nécessaire pour répondre à des niveaux de précision prédéterminés tant du point de vue de l'erreur type que du coefficient de variation des estimations. La section F donne une analyse des données de l'enquête et souligne notamment le fait qu'une analyse appropriée de ces données doit tenir compte des caractéristiques de l'échantillon dont émanent les données. La section G résume certaines questions importantes concernant la conception des enquêtes sur les ménages dans les pays en développement et les pays en transition. Un diagramme décrivant les phases importantes d'un processus d'enquête type et les liens entre les différentes phases de ce processus est présenté en annexe.

B. Échantillonnage stratifié à phases multiples

5. La plupart des enquêtes effectuées dans les pays en développement et les pays en transition sont fondées sur des échantillons en grappes stratifiés à phases multiples, et ce pour deux raisons. Premièrement, l'absence ou la mauvaise qualité des listes de ménages ou des adresses oblige d'abord à choisir un échantillon d'unités géographiques, puis à n'établir des listes des ménages ou des adresses que pour les unités choisies. Les échantillons de ménages peuvent ensuite être choisis à partir de ces listes. Deuxièmement, l'utilisation de modèles à phases multiples permet de maîtriser le coût de la collecte de données. Dans la présente section, nous examinerons les aspects statistiques et opérationnels des diverses phases d'un modèle type à phases multiples.

1. Stratification explicite

6. La stratification est couramment appliquée à chaque phase de l'échantillonnage. Cependant, ses avantages sont particulièrement visibles lors de l'échantillonnage des UPE. C'est pourquoi il importe de stratifier convenablement les UPE avant leur sélection.

7. La stratification fractionne les unités de population en sous-groupes ou strates mutuellement exclusifs ou collectivement exhaustifs. On choisit ensuite des échantillons distincts à partir de chaque strate. La stratification a premièrement pour but d'améliorer la précision des estimations de l'enquête. En l'occurrence, la composition de la strate doit être telle que les unités d'une même strate soient aussi homogènes que possibles et les unités de strates différentes aussi hétérogènes que possible pour ce qui est des caractéristiques intéressant l'enquête. La stratification présente d'autres avantages, notamment : i) la facilité et la souplesse administrative, et ii) la représentation garantie de domaines importants et de sous-groupes de population spéciaux.

8. L'expérience passée de la conception d'échantillons et de l'analyse des données dans de nombreux pays laisse apparaître de profondes différences dans la distribution des caractéristiques de la population à travers les régions administratives et entre les zones urbaines et rurales de chaque pays (voir les exemples donnés à cet égard aux chapitres XXII, XXIII et XXV ci-après). C'est là l'une des raisons pour lesquelles, pour les enquêtes effectuées dans ces pays, les strates sont généralement fondées explicitement sur les régions administratives, et sur les zones urbaines et rurales composant ces régions. Certaines de celles-ci, comme les capitales, peuvent ne pas comporter de composante rurale, tandis que d'autres peuvent être exemptes de composante urbaine. Il est recommandé d'examiner la distribution des ménages et des personnes à travers ces domaines avant de choisir telle ou telle strate pour l'échantillon.

9. Dans certains cas, des estimations sont souhaitables non seulement à l'échelon national, mais aussi séparément pour chaque région administrative ou chaque sous-région telle que province, département ou district. La stratification peut être utilisée pour contrôler la distribution de l'échantillon sur la base de ces domaines d'intérêt. Par exemple, dans les Enquêtes démographiques et sanitaires (DHS) examinées au chapitre XXII, les strates sont d'abord fondées sur les régions administratives pour lesquelles on souhaite établir des estimations. Au sein d'une même région, on procède à une nouvelle stratification entre composantes urbaines et rurales ou selon d'autres types de subdivision administrative. Des taux d'échantillonnage disproportionnés sont imposés à travers les domaines pour assurer une précision adéquate des estimations de chaque domaine. D'une façon générale, pour de nombreux domaines, la demande de données fiables nécessite l'utilisation de vastes échantillons. La question de l'estimation par domaine est examinée à la section D.

2. Stratification implicite

10. À l'intérieur de chaque strate, on utilise souvent une technique connue sous le nom de stratification implicite pour le choix des UPE. Avant de choisir l'échantillon, on trie les UPE d'une strate donnée selon une ou plusieurs variables dont on juge qu'elles ont un rapport avec la variable qui nous intéresse et qui se retrouvent dans chaque UPE de la strate en question. On choisit ensuite un échantillon systématique d'UPE. La stratification implicite garantit que les UPE de l'échantillon couvrent toutes les catégories de variables de la stratification.

11. Pour beaucoup d'enquêtes sur les ménages effectuées dans les pays en développement et les pays en transition, la stratification implicite est fondée sur la distribution géographique des unités au sein d'une strate donnée. Parmi les variables de la stratification implicite utilisées parfois pour le choix des UPE figurent le niveau de revenu (faible, moyen ou élevé), les catégories de dépenses (généralement divisées en quintiles), le groupe ethnique et le quartier habité dans les zones urbaines, et les zones cultivées, la quantité de volaille ou de bétail détenue, la proportion de travailleurs non agricoles, etc., dans les zones rurales. Pour les enquêtes socioéconomiques, les variables de la stratification implicite sont notamment la proportion de ménages classés comme pauvres, la proportion d'adultes ayant un niveau d'instruction secondaire ou supérieur, et la distance du centre dans le cas d'une grande ville. Les variables utilisées pour la stratification implicite sont généralement tirées des données des recensements.

3. Choix des UPE de l'échantillon

Caractéristiques des bonnes UPE

12. Pour les enquêtes sur les ménages effectuées dans les pays en développement et les pays en transition, les UPE sont souvent les petites unités géographiques d'une strate. Si l'on dispose des informations d'un recensement, les UPE peuvent être les zones d'énumération identifiées et utilisées pour le recensement. Des zones ou listes de population locale similaires sont parfois également utilisées. Dans les zones rurales, les villages peuvent devenir des UPE. Dans les zones urbaines, les UPE peuvent être des quartiers ou des îlots d'habitation.

13. Comme le choix des UPE influe sur la qualité de toutes les phases ultérieures du processus d'enquête, il importe de veiller à ce que les unités désignées comme UPE soient de bonne qualité et qu'elles soient choisies d'une façon raisonnablement efficace. D'une façon générale, les UPE sont jugées de bonne qualité si elles :

- a) Ont des limites identifiables et stables;
- b) Couvrent entièrement la population cible;
- c) Sont d'une dimension convenable pour l'échantillonnage;
- d) Offrent des données pour la stratification;
- e) Regroupent un nombre suffisamment élevé de ménages ou d'individus pour l'enquête.

14. Avant de choisir l'échantillon, il faut évaluer la qualité du cadre d'échantillonnage. Pour un cadre de zones d'énumération, il faut d'abord revoir les chiffres du recensement par domaine d'intérêt. En règle générale, il convient d'accorder une grande attention à la nature des UPE et à la répartition des ménages et des individus à travers les UPE pour l'ensemble de la population et pour les domaines d'intérêt. Un examen attentif de ces répartitions est nécessaire à des décisions judicieuses quant au choix des UPE et doit permettre d'identifier les unités qui ont besoin d'ajustements

pour répondre à la définition d'une bonne UPE. Le plus souvent, une grande variabilité du nombre de ménages et de personnes d'une UPE à une autre ou dans le temps nuit à la bonne organisation du travail de terrain. De même, le choix d'UPE présentant les mêmes probabilités a des effets pervers sur la précision des estimations de l'enquête.

15. Souvent, les choix naturels d'UPE ne peuvent être retenus car il manque à ces UPE une ou plusieurs des caractéristiques susmentionnées. Ces UPE ont besoin d'être modifiées ou ajustées avant de pouvoir être utilisées. Par exemple, si les limites des zones d'énumération ne paraissent pas bien définies, on peut utiliser comme UPE des unités plus vastes et plus clairement définies telles que les districts administratifs, les villages ou les communes. Par ailleurs, les UPE jugées trop vastes sont parfois fractionnées, ou peuvent être traitées comme des strates, souvent appelées UPE « autonomes » (voir Kalton, 1983). Les petites UPE sont généralement combinées avec des UPE voisines pour répondre à l'exigence d'un nombre minimum spécifié de ménages par UPE. Il est préférable d'effectuer l'ajustement des UPE trop petites ou trop grandes avant la sélection de l'échantillon.

16. Pour assurer une distribution équitable des ménages au sein des UPE, on subdivise parfois les très grandes UPE en un certain nombre de sous-unités de taille raisonnable, dont l'une est choisie au hasard pour de nouvelles opérations de terrain, telles que l'établissement d'une liste des ménages. C'est ce que l'on appelle le *morcellement* ou la *segmentation*. Il est à noter que le choix et la segmentation de trop grandes UPE introduit une phase supplémentaire dans l'échantillonnage, dont il faut tenir compte dans le processus de pondération.

17. Les UPE très petites peuvent aussi être combinées avec des UPE voisines pour répondre à l'exigence d'une taille minimum. Toutefois, on réduit considérablement le travail que comporte le regroupement de petites UPE en effectuant ce regroupement pendant ou après leur sélection. Néanmoins, c'est là un processus laborieux qui exige le respect de règles strictes et une abondante comptabilité. On trouvera une description du procédé de combinaison d'UPE pendant ou après la sélection de l'échantillon dans Kish (1965). L'inconvénient de ce procédé est qu'il ne garantit pas la contiguïté des UPE choisies pour être regroupées. Par conséquent, ce procédé n'est pas recommandé en cas de nombreuses UPE de petite taille.

Problèmes liés au manque de précision dans la mesure de la taille des UPE et solutions possibles

18. L'un des problèmes les plus courants que posent les cadres des zones d'énumération utilisées comme UPE — comme c'est souvent le cas dans les pays en développement et les pays en transition — est que la mesure de la taille peut être très imprécise. Ces mesures consistent généralement à compter les nombres de personnes ou de ménages des UPE sur la base du dernier recensement de population. Or, ces chiffres peuvent être sensiblement éloignés de la réalité actuelle, en raison de facteurs tels que la croissance des zones urbaines ou la baisse de la population d'autres régions sous l'effet de la migration, de guerres ou de catastrophes naturelles. La mesure inexacte de la taille d'une unité fait perdre le contrôle de la distribution des unités secondaires et de la taille des sous-échantillons, ce qui peut poser de sérieux problèmes lors de nouvelles opérations sur le terrain. L'une des solutions au problème du manque de précision dans la mesure de la taille des unités consiste à procéder minutieusement à l'établissement de listes afin de créer un cadre de ménages pour les UPE choisies avant de choisir les ménages. Une autre solution consiste à choisir des UPE selon une probabilité proportionnelle à la taille estimée. Ces deux procédures sont décrites de façon détaillée

aux sections 4 et 5 ci-dessous. Parmi les autres problèmes liés couramment à l'utilisation de zones d'énumération comme UPE figurent le manque de cartes de qualité et une couverture incomplète de la population cible, l'un des multiples problèmes liés au cadre d'échantillonnage examinés à la section C.

4. Échantillonnage des UPE selon une probabilité proportionnelle à leur taille

19. Avant de choisir l'échantillon, on stratifie les UPE explicitement et implicitement à l'aide de certaines des variables indiquées aux sections B.1 et B.2. Pour la plupart des enquêtes sur les ménages effectuées dans les pays en développement et les pays en transition, on choisit des UPE présentant une probabilité proportionnelle à leur taille. Avant la sélection de l'échantillon, on affecte à chaque UPE une taille généralement fondée sur le nombre de ménages ou de personnes enregistrés lors d'un recensement récent ou à la suite d'un récent exercice d'actualisation. Ensuite, on choisit un échantillon distinct d'UPE dans chaque strate explicite, présentant une probabilité proportionnelle à la taille attribuée.

20. L'échantillonnage selon une probabilité proportionnelle à la taille [*Probability proportional to size (PPS) sampling*] est une technique qui utilise des données auxiliaires pour accroître très sensiblement la précision des estimations des enquêtes, surtout si la mesure de la taille est précise et si les variables intéressantes sont liées à la taille de l'unité. C'est la méthodologie de choix pour l'échantillonnage d'UPE utilisé pour la plupart des enquêtes sur les ménages. L'échantillonnage « PPS » donne des probabilités inégales pour le choix des UPE. Par essence, la mesure de la taille de l'UPE détermine la probabilité qu'elle soit choisie. Toutefois, lorsqu'elle est combinée avec une fraction appropriée de sous-échantillon, pour le choix des ménages au sein des UPE sélectionnées, elle peut déboucher sur un échantillon global de ménages où tous les ménages présentent la même probabilité de sélection, quelle que soit l'UPE où ils se trouvent. Son principal attrait tient au fait qu'elle peut aboutir à des tailles d'échantillon à peu près égales par UPE.

21. Pour les enquêtes sur les ménages, le nombre de ménage est un bon exemple de variable PPS pour le choix des UPE. Il est admis que le nombre de ménages d'une UPE varie dans le temps et peut ne plus être d'actualité au moment de la sélection de l'échantillon. Toutefois, il y a plusieurs moyens de faire face à ce problème, comme on l'a vu au paragraphe 18. Pour les enquêtes sur les exploitations agricoles, la mesure qui est souvent retenue pour l'échantillonnage PPS est la taille de l'exploitation. Ce choix s'explique en partie par le fait que les paramètres types présentant un intérêt pour les enquêtes sur ces exploitations, tels que le revenu, la production agricole, le bétail détenu et les dépenses, sont fonction de la taille de l'exploitation. Pour les entreprises, les mesures types de la taille sont notamment le nombre d'employés, le nombre d'établissements et le volume annuel des ventes. Comme le nombre de ménages, ces mesures PPS de la taille ont des chances de varier dans le temps, et il faut tenir compte de ce fait lors de la conception de l'échantillon.

22. Considérons un échantillon de ménages, établi à partir d'un modèle à deux stades, où a UPE ont été choisies au premier stade et un échantillon de ménages ont été choisis au second stade. Supposons que la mesure de la taille (par exemple, le nombre de ménages lors du dernier recensement) de la i^{e} UPE soit M_i . Si les UPE sont choisies

selon une probabilité proportionnelle à la taille, la probabilité P_i que i^{e} UPE soit choisie est donnée par l'équation :

$$P_i = a \times \frac{M_i}{\sum_i M_i}$$

23. Maintenant, supposons que P_{ji} indique la probabilité conditionnelle que soit choisi le j^{e} ménage de la i^{e} UPE, étant donné que la i^{e} UPE a été choisie au premier stade. Dès lors, l'équation de sélection pour la probabilité inconditionnelle P_{ij} de choisir le j^{e} ménage de la i^{e} UPE est la suivante :

$$P_{ij} = P_i \times P_{ji}$$

24. Si l'on désire un échantillon de ménages de même probabilité avec une fraction d'échantillonnage global de $f = P_{ij}$, il faut choisir les ménages en nombre approprié, inversement proportionnel à la probabilité de sélection des UPE où ils se situent, c'est-à-dire :

$$P_{ji} = \frac{f}{P_i}$$

25. Si les mesures de la taille des UPE correspondent à leur taille véritable, et s'il n'y a pas de changement dans la mesure de la taille entre la sélection de l'échantillon et la collecte des données, et si b ménages sont sélectionnés dans chaque UPE de l'échantillon, on obtient un échantillon autonome de ménages dont la probabilité est donnée par l'équation :

$$P_{ij} = a \times \frac{M_i}{\sum_i M_i} \times \frac{b}{M_i} = \frac{a \times b}{\sum_i M_i} = f$$

où f est une constante.

26. Le problème que pose cette procédure est que les mesures réelles de la taille sont rarement connues dans la pratique. Toutefois, il est souvent possible d'en obtenir de bonnes estimations, telles que les comptages de la population et des ménages d'un recensement récent, ou quelque autre source fiable. Cela nous permet d'appliquer la procédure connue sous le nom d'échantillonnage selon une probabilité proportionnelle à la taille estimative [*probability-proportional-to-estimated-size (PPES) sampling*]. Il y a deux choix d'échantillonnage PPES dans un modèle à deux stades où les ménages sont sélectionnés au second stade : *a*) choisir les ménages à un taux fixe dans chaque UPE de l'échantillon; ou *b*) choisir un nombre fixe de ménages pour chaque UPE de l'échantillon.

27. L'échantillonnage PPES des ménages à raison d'un même nombre se fait de la façon suivante : appelons les valeurs vraies de la mesure de la taille N_i , et supposons que les valeurs M_i soient de bonnes estimations de N_i . Nous appliquons ensuite le taux d'échantillonnage b/M_i à la i^{e} UPE pour obtenir une taille d'échantillon de :

$$b_i = \frac{b}{M_i} \times N_i$$

28. Il est à noter que le sous-échantillonnage au sein des UPE à un taux fixe (inversement proportionnel à la mesure de la taille des UPE) suppose la détermination d'un taux pour chaque UPE de l'échantillon, de manière que, avec la probabilité de sélection d'une UPE, nous obtenions un échantillon de ménages de même probabilité, quelle que soit la taille effective des UPE. Toutefois, cette procédure ne permet pas de contrôler la taille des sous-échantillons ni, par conséquent, celle de l'échantillon. Un plus grand nombre de ménages sont choisis à partir des UPE ayant un nombre de ménages plus élevé que prévu, et vice versa. Cela présente des conséquences pour l'organisation du travail de terrain. De plus, si les mesures de la taille sont si dépassées que les différences entre échantillons sont extrêmes, il peut se révéler nécessaire de modifier le taux d'échantillonnage de manière à obtenir des tailles d'échantillon légèrement plus homogènes entre les UPE, ce qui impliquerait que l'on s'écarte du modèle à auto-pondération.

29. La seconde procédure, la sélection d'un nombre fixe de ménages par UPE, permet d'éviter l'inconvénient de tailles d'échantillon variables selon les UPE, mais ne permet pas de produire un échantillon à auto-pondération. Toutefois, si les mesures de la taille sont actualisées immédiatement avant la sélection de l'échantillon d'UPE, elles peuvent donner des approximations suffisamment bonnes pouvant déboucher sur un échantillon de ménages plus ou moins autopondéré.

30. En résumé, bien que le sous-échantillonnage d'UPE à taux fixe soit conçu pour produire des échantillons autopondérés, il y a des cas où cette méthode oblige à s'écarter de ce type d'échantillons. En revanche, s'il est fréquent que le choix d'un nombre fixe de ménages au sein des UPE ne produise pas d'échantillons autopondérés, dans certains cas, cette méthode débouche sur des échantillons de ménages plus ou moins autopondérés. Chaque fois que l'on s'écarte du modèle autopondéré, il faut utiliser les pondérations pour compenser les différences de probabilités de sélection entre différentes UPE.

5. Sélection d'un échantillon de ménages

31. Une fois que la sélection des UPE de l'échantillon est terminée, on procède à l'établissement de la liste de tous les ménages ou de tous les logements ou habitations de chaque UPE sélectionnée. Parfois, on inclut la liste des unités d'habitation, puis tous les ménages des unités d'habitation choisies sont inclus si une unité d'habitation est retenue pour l'échantillon. Le but de ces listes est de créer un cadre d'échantillonnage à jour à partir duquel on puisse choisir les ménages. On ne saurait trop insister sur l'importance de la bonne exécution de cette opération. La qualité de l'établissement de ces listes est l'un des facteurs dont l'influence sur la couverture de la population cible est l'une des plus importantes.

32. Avant de procéder à la sélection de chaque UPE de l'échantillon, on peut classer les ménages figurant sur les listes par rapport à des variables géographiques ou autres considérées comme ayant une relation étroite avec les variables intéressant l'enquête (voir section B.2). Ensuite, des ménages sont choisis sur la liste selon une procédure d'échantillonnage systématique à probabilité égale. Comme il est indiqué à la section B.4, les ménages peuvent être sélectionnés à l'intérieur des UPE de l'échantillon à des taux d'échantillonnage qui engendrent des probabilités globales de sélection égales pour tous les ménages ou à des taux qui engendrent un nombre fixe de ménages choisis dans chaque UPE. Les avantages et inconvénients de ces approches sont examinés à la section B.4.

33. Souvent, les unités d'échantillonnage ultimes sont les ménages, et les informations sont recueillies sur les ménages sélectionnés et sur tous les membres de ces

ménages. Pour les modules spéciaux portant sur les revenus et les dépenses, pour lesquels les ménages sont les unités d'analyse, on choisit souvent un membre averti pour fournir les réponses sur le ménage. Pour les questions jugées délicates pour certains membres de la famille (par exemple, les mauvais traitements), on choisit un échantillon aléatoire (souvent composé d'une seule personne) au sein de chaque ménage sélectionné.

6. Nombre de ménages à choisir par UPE

34. Les unités primaires d'échantillonnage sont constituées d'un ensemble de ménages concentrés en un même lieu géographique. Ainsi, les ménages d'un même groupe ont généralement tendance à présenter des caractéristiques plus homogènes du point de vue de l'enquête (par exemple, de revenu, d'instruction, de profession, etc.) que les ménages en général. Le regroupement en grappes permet de réduire considérablement le coût de collecte des données, mais les corrélations entre unités d'une même grappe accroissent la variance (réduisent la précision) des estimations de l'enquête, par comparaison avec une conception où les ménages ne sont pas regroupés en grappes. Autrement dit, la difficulté pour celui qui conçoit l'enquête est de parvenir au bon équilibre entre les économies réalisées et la perte de précision liée à la mise en grappes.

35. L'accroissement de la variance des estimations de l'enquête imputable au regroupement en grappes contribue à ce que l'on appelle l'effet lié à la conception. Cet effet représente le coefficient par lequel la variance d'une estimation fondée sur un simple échantillon aléatoire de même taille doit être multipliée pour tenir compte des complexités de la conception de l'échantillon dues à la stratification, à la formation en grappes et à la pondération. Il est défini comme le ratio de la variance d'une estimation fondée sur la conception complexe à celle fondée sur un simple échantillon aléatoire de même taille. Voir les chapitres VI et VII de la présente publication et les références qui y sont citées pour plus de détails sur les effets liés à la conception et leur utilisation dans la conception de l'échantillon. Une expression de l'effet lié à la conception (dû à la formation de grappes) pour une estimation [par exemple, une moyenne estimatives (\bar{y})] est donnée approximativement par l'équation :

$$D^2(\bar{y}) = 1 + (b - 1)\rho$$

où $D^2(\bar{y})$ indique l'effet lié à la conception pour la moyenne estimative (\bar{y}), ρ est la corrélation au sein d'un même groupe, et b est le nombre moyen de ménages à sélectionner au sein de chaque grappe, c'est-à-dire la taille moyenne d'un échantillon en grappes. La corrélation au sein d'un même groupe mesure le degré d'homogénéité (par rapport à la variable intéressante) des unités d'une même grappe. Comme les unités d'une même grappe tendent à être similaires les unes par rapport aux autres, la corrélation au sein d'un même groupe est presque toujours positive. Pour les populations, une corrélation positive au sein d'un même groupe peut être due au fait que les ménages d'une même grappe appartiennent à la même catégorie de revenus, partagent les mêmes attitudes envers les problèmes du jour et sont souvent exposés aux mêmes conditions d'environnement (climat, maladies infectieuses, catastrophes naturelles, etc.).

36. Le fait de ne pas tenir compte de l'effet lié à la conception dans les estimations des erreurs types peut entraîner une interprétation erronée des résultats de l'enquête. Il convient de noter que l'ampleur de $D^2(\bar{y})$ est liée directement à la valeur de b , à la taille de l'échantillon en grappes et à la corrélation au sein d'un même

groupe (ρ). Pour une valeur fixe de ρ , l'effet lié à la conception augmente de façon linéaire avec b . Ainsi, pour que les effets liés à la conception soient faibles, il est préférable d'utiliser un échantillon en grappes aussi petit que possible. Le tableau II.1 illustre la façon dont la taille moyenne de l'échantillon en grappes et la corrélation au sein d'un même groupe influent sur l'effet lié à la conception. Par exemple, avec un échantillon en grappes moyen b de 20 unités d'habitation par UPE et ρ égal à 0,05, l'effet lié à la conception est de 1,95. Autrement dit, cette conception de l'échantillon en grappes donne des estimations présentant la même variance que celles d'un échantillon non en grappes (échantillon aléatoire simple) d'environ la moitié du nombre total de ménages. Avec des valeurs de ρ plus élevées, la perte de précision est encore plus forte, comme le montrent les chiffres des colonnes de droite du tableau II.1.

Tableau II.1
Effets liés à la conception pour certaines combinaisons de taille d'échantillon en grappes et de corrélation au sein d'un même groupe

Taille de l'échantillon en grappes (b)	Corrélation au sein d'un même groupe (ρ)								
	0,005	0,01	0,02	0,03	0,04	0,05	0,10	0,20	0,30
1	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
10	1,05	1,09	1,18	1,27	1,36	1,45	1,90	2,80	3,70
15	1,07	1,14	1,28	1,42	1,56	1,70	2,40	3,80	5,20
20	1,10	1,19	1,38	1,57	1,76	1,95	2,90	4,80	6,70
30	1,15	1,29	1,58	1,87	2,16	2,45	3,90	6,80	9,70
50	1,25	1,49	1,98	2,47	2,96	3,45	5,90	10,80	15,70

37. D'une façon générale, le nombre optimum de ménages à choisir au sein de chaque UPE dépend de la structure des coûts de la collecte de données et du degré d'homogénéité ou de la formation de grappes par rapport aux variables de l'enquête au sein de l'UPE. Supposons une conception en deux phases, où les UPE sont choisies dans la première phase et les ménages dans la seconde. Supposons également que le modèle de coût linéaire pour l'ensemble des coûts liés à l'échantillonnage des UPE et des ménages soit donné par l'équation :

$$C = aC_1 + abC_2$$

où C_1 et C_2 sont respectivement le coût d'une UPE supplémentaire et le coût d'un ménage supplémentaire, et a et b indiquent respectivement le nombre d'UPE sélectionnées et le nombre de ménages sélectionnés pour chaque UPE (Cochran, 1977, p. 280). Selon ce modèle de coût, le choix optimum pour b qui minimise la variance de l'échantillon moyen (voir Kish, 1965, sect. 8.3.b) est donné approximativement par :

$$b_{opt} = \sqrt{\frac{C_1(1-\rho)}{C_2 \rho}}$$

38. Le tableau II.2 donne la taille optimale du sous-échantillon (b) pour divers ratios C_1/C_2 et corrélations au sein d'un même groupe. On notera que, toutes choses égales par ailleurs, la taille optimale de l'échantillon diminue (c'est-à-dire que l'échantillon est prélevé plus largement à travers les grappes) à mesure que la corrélation au sein d'un même groupe augmente, et à mesure que le coût d'un ménage additionnel augmente par rapport à celui de l'UPE.

39. Le modèle de coût utilisé dans la détermination de la taille optimal d'une grappe est très simplifié, mais probablement adéquat comme indication générale. Comme, par nature, la plupart des enquêtes ont plusieurs buts, nécessitant des variables différentes et des valeurs différentes de ρ , le choix de b suppose souvent un certain degré de compromis entre plusieurs optimums différents.

Tableau II.2.

Tailles optimales de sous-échantillons pour certaines combinaisons de ratio de coût et de corrélation au sein d'un même groupe

Ratio de coût (C_1/C_2)	Corrélation au sein d'un même groupe				
	0,01	0,02	0,03	0,05	0,08
4	20	14	11	9	5
9	30	21	17	13	10
16	40	28	23	17	14
25	50	35	28	22	17

40. En l'absence de données précises sur les coûts, le tableau II.2 peut être utilisé pour déterminer le nombre optimal de ménages à choisir dans une grappe pour divers choix de ratio de coût et de corrélation au sein d'un même groupe. Par exemple, si l'on sait a priori que le coût d'inclusion d'une UPE est quatre fois plus élevé que le coût d'inclusion d'un ménage, et que la corrélation au sein d'un même groupe pour une variable à étudier est de 0,05, il est alors judicieux de choisir environ neuf ménages de la grappe. Il convient de noter que le nombre optimal de ménages à choisir dans une grappe n'est pas fonction du budget global de l'enquête. Le budget total ne permet de déterminer que le nombre d'UPE à choisir.

41. En règle générale, les facteurs à prendre en compte lors de la détermination du choix des UPE et des ménages à inclure dans l'échantillon sont notamment la précision des estimations de l'enquête (à travers l'effet lié à la conception), le coût de collecte des données et l'organisation du travail de terrain. Si les coûts de déplacement sont élevés, comme dans le cas des zones rurales, il est préférable de ne choisir qu'un petit nombre d'UPE mais un grand nombre de ménages dans chaque UPE. En revanche, comme dans les zones urbaines, les coûts de déplacement sont faibles, et il est préférable de choisir de nombreuses UPE et moins de ménages par UPE. Ces choix doivent tendre à assurer une distribution efficace de la charge de travail entre les enquêteurs et les superviseurs.

C. Cadres d'échantillonnage

1. Caractéristiques des cadres d'échantillonnage retenus pour les enquêtes dans les pays en développement et les pays en transition

42. Pour la plupart des enquêtes sur les ménages, la population cible est composée de la population civile ne vivant pas dans une institution. Pour obtenir les données attendues de cette population cible, les entrevues se déroulent souvent au sein du ménage. En général, seules les personnes considérées comme membres permanents du ménage sont habilitées à être incluses dans les enquêtes. Les membres permanents d'un ménage qui sont temporairement absents, tels que les personnes en vacances, ou temporairement hospitalisées, et les étudiants vivant loin de leur foyer pendant l'année scolaire, sont généralement inclus si le ménage est choisi. Les étudiants vivant loin de

leur foyer pendant l'année scolaire ne sont pas inclus dans l'enquête s'ils sont retenus pour l'échantillon à l'endroit où ils font leurs études car les données les concernant seraient déjà obtenues à leur lieu de résidence permanente. Parmi les personnes qui sont généralement exclues des enquêtes sur les ménages dans les pays en développement et les pays en transition figurent les membres des forces armées vivant dans des casernes ou des logements particuliers, les personnes se trouvant en prison, à l'hôpital, dans des maisons de santé ou de retraite ou dans d'autres institutions, les sans-logis et les nomades. La plupart de ces groupes sont généralement exclus en raison des difficultés pratiques que comporte habituellement la collecte de données auprès d'eux. Toutefois, la décision d'exclure ou non un groupe doit être fonction des objectifs de l'enquête.

2. Problèmes liés aux cadres d'échantillonnage et solutions possibles

43. Comme pour les autres types d'enquêtes, la qualité des données recueillies auprès des ménages dépend pour beaucoup de la qualité du cadre d'échantillonnage à partir duquel a été choisi l'échantillon retenu pour l'enquête. Malheureusement, les problèmes liés aux cadres d'échantillonnage sont l'un des traits inévitables des enquêtes sur les ménages. La présente section examine certains de ces problèmes et leur suggère des solutions possibles.

44. Kish (1965, sect. 2.7) offre une classification utile de quatre de ces problèmes et les solutions à leur apporter. Ces quatre problèmes sont l'absence de couverture, les grappes d'éléments, les blancs et les doubles inscriptions sur les listes.

45. L'expression « absence de couverture » désigne le fait que le cadre d'échantillonnage ne couvre pas la totalité de la population cible, c'est-à-dire que certaines unités d'échantillonnage n'ont aucune probabilité d'inclusion dans l'échantillon. L'absence de couverture est l'une des grandes préoccupations qu'inspirent les enquêtes sur les ménages dans les pays en développement et les pays en transition. Son incidence se manifeste par le fait que la qualité des estimations des comptages de population fondées sur la plupart des enquêtes effectuées dans les pays en développement et les pays en transition n'a pas été à la hauteur de celles des estimations de populations obtenues d'autres sources.

46. L'absence de couverture s'observe à trois niveaux : au niveau des UPE, au niveau des ménages et au niveau des individus. Pour les pays en développement et les pays en transition, l'absence de couverture des UPE est moins grave que celle des ménages et des personnes retenues au niveau des ménages sélectionnés pour l'échantillon. L'absence de couverture des UPE se produit notamment lorsque certaines régions d'un pays sont exclues à dessein d'une enquête, parce qu'elles sont inaccessibles, en guerre, parce qu'elles sont le théâtre d'une catastrophe naturelle ou pour d'autres raisons. Par ailleurs, les régions reculées où vivent peu de ménages ou d'individus sont parfois retirées des cadres d'échantillonnage des enquêtes sur les ménages parce qu'elles ne représentent qu'une faible proportion de la population ou parce qu'elles n'ont guère d'effets sur les chiffres de population. L'absence de couverture pose un problème plus grave aux niveaux des ménages et des personnes. Des ménages ou des personnes peuvent se trouver à tort exclus de l'enquête par suite de problèmes complexes de définition ou de concept concernant la structure et la composition des ménages. On court le risque d'incohérences dans l'interprétation de ces questions par les différents enquêteurs ou les personnes chargées de dresser les listes des ménages et de leurs membres. C'est pourquoi on a besoin d'instructions opérationnelles rigoureuses pour guider les enquêteurs sur les questions de savoir qui doit être considéré comme membre d'une famille ou ce qui doit être considéré comme un ménage ou une unité d'habitation.

Pour faire face à ce problème, il conviendrait de faire de la qualité des listes de ménages et de personnes admissibles au sein de ces ménages une question clé du travail méthodologique et de la formation dans les pays en développement et les pays en transition.

47. Le problème des blancs se pose lorsque certaines listes du cadre d'échantillonnage ne contiennent pas d'éléments de la population cible. Dans le cas des listes d'unités d'habitation, le blanc correspond à un logement vide. Ce problème se pose également si l'on effectue un échantillonnage sur certains sous-groupes particuliers, par exemple les femmes ayant eu un enfant l'année précédente. Certains ménages inscrits sur la liste et choisis pour l'échantillon n'ont aucune femme ayant eu un enfant l'année précédente. Dans la mesure du possible, les blancs doivent être retirés du cadre avant la sélection de l'échantillon. Toutefois, dans bien des applications pratiques, cela n'est pas économique. Il est préférable d'identifier et d'éliminer les blancs après la sélection de l'échantillon. Cela dit, l'élimination des blancs a pour effet de produire des échantillons plus petits et de taille inégale.

48. Le problème des doubles inscriptions sur les listes se produit lorsque des unités de la population cible apparaissent plus d'une fois dans le cadre d'échantillonnage. Tel est le cas, par exemple, lorsque des nomades ou des résidents à temps partiel font partie de l'échantillon. Un moyen d'éviter les doubles inscriptions est de désigner une liste unique prédéterminée comme la liste effective et les autres comme blancs. L'unité n'est incluse dans l'échantillon que si la liste unique est retenue. Par exemple, les nomades qui déplacent leurs troupeaux d'un endroit à un autre à la recherche de pâturages et d'eau pour leurs bêtes peuvent être enregistrés lorsqu'ils se rendent à un point d'eau. Selon les cycles d'abreuvement des bêtes (les chevaux s'abreuvent moins souvent que les bovins), certains ont des chances de visiter plus d'un point d'eau pendant la période de collecte de données pour l'enquête. Pour éviter les doubles inscriptions, on pourrait n'enregistrer les nomades qu'à leur première visite à un point d'eau après une certaine date et traiter les visites ultérieures comme des blancs. Sinon, il faudrait ajuster les pondérations des unités de l'échantillon pour tenir compte de ces doubles inscriptions. Voir les exemples de la façon de procéder donnés par Yansaneh (2003).

49. Le problème des grappes d'éléments se pose lorsqu'une liste unique du cadre d'échantillonnage comprend de multiples unités de la population cible. Par exemple, une liste de logements peut comprendre des logements où habite plus d'un ménage. En pareils cas, l'inclusion de tous les ménages liés au logement de l'échantillon produit un échantillon où les ménages ont la même probabilité d'être sélectionnés que le logement. Il convient de noter que la pratique consistant à choisir au hasard l'une des unités d'une grappe conduit automatiquement à des probabilités inégales de sélection, qu'il conviendrait alors de compenser par une pondération.

3. Maintien et évaluation des cadres d'échantillonnage

50. La construction et le maintien de bons cadres d'échantillonnage est une opération coûteuse et de longue haleine. Les pays en développement et les pays en transition ont la possibilité de créer ces cadres à partir de sources telles que les données de leurs recensements décennaux. Il est recommandé que chaque office national de statistique se fixe comme grande priorité la création et le maintien d'un cadre directeur d'échantillonnage des zones d'énumération qui ont été définies et utilisées lors d'un précédent recensement. Ce cadre devrait être créé peu après l'achèvement du recensement, car le volume de travail nécessaire augmente à mesure que l'on s'éloigne de la période du recensement. Ce cadre doit étiqueter clairement les autres zones géographiques, parfois plus vastes, pouvant servir d'unités primaires d'échantillonnage. Il doit aussi comprendre des données qui peuvent être utiles pour la stratification, telles

que la composition ethnique ou raciale, le niveau médian de dépenses ou les niveaux de dépenses par quintiles, etc. S'il est tenu convenablement, le cadre directeur d'échantillonnage peut être utilisé pour la mise en œuvre d'un système intégré d'enquêtes, y compris d'enquêtes répétées. Pour plus de détails sur la construction et le maintien du cadre directeur d'échantillonnage, voir chapitre V.

D. Estimation des domaines

1. Nécessité d'estimations de domaines

51. Ces dernières années, dans la plupart des pays, il s'est produit un accroissement de la demande de données fiables non seulement à l'échelon national, mais aussi aux niveaux sous-nationaux ou domaines, dû principalement au fait que la plupart des programmes de développement ou d'intervention se déroulent aux niveaux sous-nationaux, tels qu'à ceux des régions administratives ou des districts. Pour prendre des décisions importantes concernant l'exécution de programme ou l'allocation des ressources à l'échelon local, il faut disposer de données précises à ce niveau.

52. Aux fins de la présente analyse, nous définirons un domaine comme tout sous-ensemble de population pour lequel des estimations distinctes sont envisagées dans le cadre de l'enquête. Un domaine peut être une strate, une combinaison de strates, une région administrative ou une subdivision urbaine, rurale ou autre au sein d'une telle région. Ainsi, des estimations tirées de nombreuses enquêtes nationales sont publiées séparément pour les régions administratives. Ces régions peuvent alors être traitées comme des domaines, comptant chacun deux strates (par exemple, les sous-populations urbaines et rurales) ou plus. Les domaines peuvent également être des sous-groupes démographiques définis par certaines caractéristiques telles que l'âge, la race et le sexe. Toutefois, on se heurte à une complication lorsque les domaines franchissent les limites d'une strate, comme, par exemple, dans le cas où un domaine est composé de ménages ayant accès à des services de santé.

53. Il importe que le nombre de domaines sur lesquels porte une enquête particulière soit limité. L'échantillon requis pour l'établissement d'estimations fiables pour chacun des domaines d'un grand nombre d'entre eux devrait nécessairement être très grand. Les problèmes liés aux gros échantillons seront examinés à la section E.

2. Allocation d'échantillons

54. L'établissement d'estimations précises pour les domaines faisant l'objet de l'enquête exige que des échantillons de taille adéquate soient alloués à ces domaines. Or, on se heurte à des conflits lorsque la même précision est demandée à des domaines aux chiffres de population très différents. Si l'on veut que les estimations présentent le même niveau de précision pour tous les domaines, le moyen le plus efficace est l'égalité d'allocation (c'est-à-dire un échantillon de même taille pour chaque domaine). Toutefois, une telle allocation peut occasionner une sérieuse perte d'efficacité pour les estimations nationales. L'allocation proportionnelle, qui applique des fractions d'échantillonnage égales à chaque domaine, est souvent la meilleure solution pour les estimations nationales. Lorsque les domaines sont de tailles très différentes et lorsque l'on a besoin d'estimations nationales et par domaine, il faut trouver un compromis entre l'allocation à égalité et des fractions d'échantillonnage égales.

55. Un compromis entre l'allocation proportionnelle et l'allocation égale a été proposé par Kish (1988), sur la base d'une allocation proportionnelle à $n\sqrt{(W_h^2 + H^{-2})}$ où n est la taille de l'échantillon global, W_h la proportion de la population de la strate h ,

et H le nombre de strates. Pour les très petites strates, le second terme domine le premier, empêchant ainsi les allocations aux strates trop petites.

56. L'autre option serait d'accroître la taille de l'échantillon des petits domaines de manière à répondre aux niveaux de précision requis. Lorsqu'un domaine est petit, l'allocation proportionnelle donnerait un échantillon trop petit pour permettre des estimations suffisamment précises. La solution est de créer un échantillon de plus grande taille, ou représentant une plus forte proportion, pour les petits domaines.

57. En résumé, dans les pays en développement et les pays en transition, les responsables de la conception des enquêtes ont souvent à choisir entre des estimations précises à l'échelon national et des estimations précises pour les domaines. Ce problème s'aggrave lorsque les domaines sur lesquels porte l'enquête sont de tailles très différentes. Le moyen d'éviter ce dilemme est de définir des domaines de taille plus ou moins égale, éventuellement en regroupant des domaines existants. Inversement, on peut garder distincts ces domaines et autoriser un plus faible niveau de précision pour les petits domaines, voire s'abstenir de publier des estimations pour ceux-ci.

E. Taille des échantillons

1. Facteurs qui influent sur les décisions concernant la taille des échantillons

58. Les producteurs comme les utilisateurs de données d'enquêtes désirent souvent des échantillons de grande taille, parce qu'ils jugent cela nécessaire pour rendre l'échantillon plus « représentatif », et pour réduire l'erreur d'échantillonnage et accroître ainsi la fiabilité des estimations de l'enquête. Cet argument est avancé pratiquement sans qu'il soit tenu compte de l'accroissement possible des erreurs non liées à l'échantillonnage, qui découlent de l'utilisation d'échantillons de grande taille. Dans la présente section, nous examinerons les facteurs dont il faut tenir compte lorsque l'on détermine la taille appropriée à donner à l'échantillon.

59. Les trois grandes questions dont dépendent les décisions concernant la taille appropriée de l'échantillon à établir pour une enquête sont :

- La précision (fiabilité) des estimations de l'enquête;
- La qualité des données recueillies dans le cadre de l'enquête;
- Le coût en temps et en argent de la collecte, du traitement et de la diffusion des données.

Nous examinerons à présent chacun de ces facteurs à tour de rôle.

2. Précision des estimations de l'enquête

60. La plupart des enquêtes effectuées dans les pays en développement et les pays en transition ont pour objet l'estimation du niveau d'une caractéristique (par exemple, le pourcentage de ménages classés comme pauvres) à un moment déterminé et le pourcentage de variation de ce niveau dans le temps (par exemple, la variation du taux de pauvreté entre deux dates). Nous examinerons la précision des estimations de l'enquête dans le contexte de l'estimation du niveau d'une caractéristique à un moment déterminé.

61. La précision d'une estimation se mesure à son erreur type. La formule d'estimation de l'erreur type d'un taux de pauvreté estimatif p dans un domaine donné, indiquée par les lettres $se(p)$, est donnée par l'équation :

$$se(p) = \sqrt{d^2(p) \times \left(1 - \frac{n}{N}\right) \times \frac{p(100-p)}{n}}$$

² Bien que n soit en réalité $n - 1$ dans la formule ci-dessus pour $se(p)$, dans la plupart des applications pratiques, n est suffisamment élevé pour que la différence entre n et $n - 1$ soit négligeable

où n est le nombre total de ménages du domaine étudié, N le total de ménages du domaine examiné et $d^2(p)$ l'effet estimatif lié à la complexité de conception de l'enquête². La proportion de la population de l'échantillon, n/N , est appelée la fraction d'échantillonnage, et le facteur $[1 - (n/N)]$ (la proportion de la population non comprise dans l'échantillon), est appelé le facteur de correction de la population finie (*fpc*). Le *fpc* représente l'ajustement apporté à l'erreur type pour tenir compte du fait que l'échantillon est sélectionné sans remplacement à partir d'une population finie.

Tableau II.3

Erreurs types et intervalles de confiance pour des estimations du taux de pauvreté fondées sur diverses tailles d'échantillon, avec un effet lié à la conception supposé de 2,0

Taux de pauvreté (en pourcentage)										
	5		10		25		40		50	
Taille de l'échantillon	Erreur type	Intervalle de confiance								
250	1,95	(1,2, 8,8)	2,68	(4,7, 15,3)	3,87	(17,4, 32,6)	4,38	(31,4, 48,6)	4,47	(41,2, 58,8)
500	1,38	(2,3, 7,7)	1,90	(6,3, 13,7)	2,74	(19,6, 30,4)	3,10	(33,9, 46,1)	3,16	(43,8, 56,2)
750	1,13	(2,8, 7,2)	1,55	(7,0, 13,0)	2,24	(20,6, 29,4)	2,53	(35,0, 45,0)	2,58	(44,9, 55,1)
1 000	0,97	(3,1, 6,9)	1,34	(7,4, 12,6)	1,94	(21,2, 28,8)	2,19	(35,7, 44,3)	2,24	(45,6, 54,4)
1 500	0,80	(3,4, 6,6)	1,10	(7,9, 12,1)	1,58	(21,9, 28,1)	1,79	(36,5, 43,5)	1,83	(46,4, 53,6)
2 000	0,44	(4,1, 5,9)	0,95	(8,1, 11,9)	1,37	(22,3, 27,7)	1,55	(37,0, 43,0)	1,58	(46,9, 53,1)

62. Nous utiliseront les données du Viet Nam à titre d'illustration. Le nombre total de ménages, N , d'après le recensement de la population de 1999, est de 16 661 366. Voir Glewwe and Yansaneh (2000) pour plus de détails sur la distribution des ménages fondée sur le recensement de 1999. Il convient de noter qu'avec un chiffre de population d'une telle ampleur, le facteur de correction de la population finie est négligeable dans tous les cas. Le tableau II.3 indique des erreurs types et des intervalles de confiance de 95 % pour diverses estimations du taux de pauvreté, en supposant un effet lié à la conception de 2,0. Un intervalle de confiance de 95 % signifie que la probabilité que l'on ait la valeur vraie est de 95 %. Le tableau montre que, pour un échantillon d'une taille donnée, les erreurs types augmentent à mesure qu'augmente le taux de pauvreté, atteignant un maximum pour $p = 50$ %. Les intervalles de confiance correspondants de 95 % s'élargissent aussi à mesure qu'augmente le taux de pauvreté, atteignant leur maximum lorsque ce taux est de 50 %. Ainsi, en général, pour les domaines où le taux de pauvreté est nettement inférieur ou nettement supérieur à 50 %, les estimations de l'enquête sont plus précises que pour les domaines où le taux de pauvreté approche les 50 %, pour des chiffres donnés concernant la taille de l'échantillon et l'effet lié à la conception³. Cela signifie que les domaines où le taux de pau-

vreté est très faible ou très élevé nécessitent un échantillon plus petit pour produire la même erreur type qu'un domaine où le taux de pauvreté est proche de 50 %. Examinons, par exemple, un échantillon de 500 ménages d'un domaine. Si ce domaine ne présente qu'un taux de pauvreté de 5 %, l'intervalle de confiance est de $5 \pm 2,7$ %; si ce domaine a un taux estimatif de pauvreté de 10 %, l'intervalle de confiance est de $10 \pm 3,7$ %; si ce domaine a un taux estimatif de pauvreté de 25 %, l'intervalle de confiance est de $25 \pm 5,4$ %; et si ce domaine a un taux estimatif de pauvreté de 50 %, l'intervalle de confiance est de $50 \pm 6,2$ %.

63. Bien entendu, en augmentant la taille de l'échantillon à plus de 500 ménages, on réduit l'ouverture de l'intervalle de confiance (autrement dit, l'estimation de l'échantillon devient plus précise). Toutefois, cette diminution est proportionnelle non pas à l'accroissement de la taille de l'échantillon, mais à la racine carrée de cet accroissement, en l'occurrence $\sqrt{n/500}$ où n est la taille du nouvel échantillon. Par exemple, dans un domaine où le taux de pauvreté est de 25 %, en doublant la taille de l'échantillon de 500 à 1 000 ménages, on réduit l'ouverture de l'intervalle de confiance selon un facteur de $\sqrt{2}$, c'est-à-dire de $\pm 5,4$ % à $\pm 3,8$ %. Ces réductions sont à considérer en regard de la complexité accrue de gestion de l'enquête, des coûts de l'enquête et des erreurs autres non dues à l'échantillonnage.

64. La précision des estimations de l'enquête est souvent définie par le coefficient de variation de l'estimation de l'intérêt. Comme auparavant, nous limiterons notre attention à l'estimation du pourcentage de ménages classés comme pauvres dans le pays. Le coefficient estimatif de variation d'une estimation du taux de pauvreté, indiqué par $cv(p)$, est donné par l'équation :

$$cv(p) = \frac{se(p)}{p} = \sqrt{d^2(p) \times \left(1 - \frac{n}{N}\right) \times \frac{(100-p)}{np}}$$

³ Pour les taux de pauvreté supérieurs à 50 % ($p > 50$ %), l'erreur type est la même que pour un taux de pauvreté de $100 - p$, et peut donc être déterminée à partir du tableau III.3. Par exemple, l'erreur type d'un taux estimatif de pauvreté de 75 % est la même que celle d'un taux estimatif de pauvreté de 25 %.

Tableau II.4

Coefficient de variation des estimations du taux de pauvreté fondé sur diverses tailles d'échantillon, l'effet supposé lié à la conception étant de 2,0

Taille de l'échantillon	Taux de pauvreté (en pourcentage)								
	5	10	25	40	50	60	75	90	95
250	39	27	15	11	9	7	5	3	2
500	28	19	11	8	6	5	4	2	1
750	23	15	9	6	5	4	3	2	1
1 000	19	13	8	5	4	4	3	1	1
1 500	16	11	6	4	4	3	2	1	1
2 000	14	9	5	4	3	3	2	1	1

65. Le tableau II.4 présente les coefficients estimatifs de variation d'un taux estimé de pauvreté pour diverses tailles d'échantillon, avec un effet supposé lié à la conception de 2,0, où cv est exprimé en pourcentage. Ce tableau montre que, pour un échantillon d'une taille donnée, le coefficient estimatif de variation du taux estimé de pauvreté diminue régulièrement à mesure que le pourcentage vrai augmente. En outre, pour un taux de pauvreté donné, le coefficient de variation diminue à mesure que diminue la taille de l'échantillon. Pour un échantillon de 500, le coefficient de variation est d'environ 28 % lorsque $p = 5$ %, 19 % lorsque $p = 10$ %, 11 % lorsque $p = 25$ %, 8 % lorsque $p = 40$ %, 6 % lorsque $p = 50$ %, 5 % lorsque $p = 60$ %, 4 % lorsque $p = 75$ %, 2 % lorsque $p = 90$ %, et 1 % lorsque $p = 95$ %. À mesure que l'on augmente la taille

de l'échantillon, le coefficient estimatif de variation diminue de façon correspondante. On notera qu'à la différence des erreurs types indiquées au tableau II.3, le coefficient de variation indiqué au tableau II.4 n'est pas une fonction symétrique du taux de pauvreté.

3. Qualité des données

66. La qualité des données recueillies est un important facteur dans la détermination de la taille de l'échantillon. Il est important de maintenir des données de la meilleure qualité possible, afin que l'on puisse avoir confiance dans les estimations qui en sont tirées. Il est essentiel de vérifier la qualité des données à chaque stade de l'exécution de l'enquête. Il est donc important de maintenir la taille de l'échantillon dans des limites raisonnables, afin que l'on puisse procéder à des vérifications et à des modifications d'une façon efficace à la fois en temps et en coût.

67. Le nombre de personnes travaillant à l'étude est également fonction de la taille de l'échantillon et influe sur la qualité des données. Par exemple, des échantillons plus petits nécessitent moins d'enquêteurs, de sorte que ceux-ci peuvent être choisis selon des critères plus rigoureux. En effet, avec un échantillon plus petit, il y a plus de chances que tous les enquêteurs soient recrutés parmi un personnel qualifié et expérimenté. De plus, les enquêteurs sont mieux préparés car lorsqu'ils sont moins nombreux, leur formation peut être mieux focalisée et peut se voir affecter proportionnellement plus de ressources sur le budget de l'enquête. Il faut moins de matériel didactique, et les enquêteurs reçoivent une attention plus personnalisée pendant leur formation et sur le terrain. Tout cela permet de réduire les problèmes liés à la collecte des données et à l'édition des données recueillies. Autrement dit, les données disponibles pour l'analyse sont de meilleure qualité, permettant aux décideurs de prendre leurs décisions avec plus de confiance.

68. Outre les préoccupations concernant la qualité des données recueillies, de plus gros échantillons accroissent la difficulté et le coût de réduction du taux de non-réponse à l'enquête (voir chapitre VIII). Il importe de maintenir ce taux aussi faible que possible, afin de réduire la possibilité de fortes distorsions dans les estimations de l'enquête (voir section F.1). De telles distorsions sont possibles si nous n'obtenons pas de réponses d'un fort pourcentage de personnes pouvant être sensiblement différentes de celles incluses dans l'enquête. Par exemple, les personnes qui vivent dans les zones urbaines et ont des revenus relativement élevés sont souvent moins enclines à participer aux enquêtes sur les ménages. L'exclusion d'un large segment de cette fraction de la population peut entraîner une sous-estimation de certaines caractéristiques de cette population, telles que le revenu moyen des ménages, le niveau d'instruction et le niveau d'alphabétisation à l'échelon national. Avec un échantillon plus petit, il est beaucoup plus facile et plus économique de rendre à nouveau visite aux ménages qui ont choisi au départ de ne pas participer à l'enquête, afin de les persuader de revenir sur leur décision. Comme il peut être coûteux et laborieux de persuader ces non-participants initiaux, il importe, pour la qualité des données de l'enquête, de donner aux meilleurs enquêteurs assez de temps et de ressources pour qu'ils puissent les faire changer d'avis.

4. Coût et opportunité

69. La taille de l'échantillon a certainement une influence sur le coût de l'enquête. D'une façon générale, le coût global d'une enquête est fonction de frais généraux fixes et des coûts variables liés à la sélection et au traitement de chaque unité

d'échantillonnage à chaque stade de la sélection de l'échantillon. Par conséquent, plus l'échantillon est grand, plus le coût d'exécution de l'enquête est élevé. On trouvera un examen plus détaillé des composantes pertinentes du coût des enquêtes sur les ménages au chapitre XII, et des exemples empiriques de la détermination des coûts de certaines enquêtes spécifiques aux chapitres XIII et XIV.

70. La taille de l'échantillon peut également influencer sur le moment où les données sont fournies pour l'analyse. Il importe que les données et les estimations de l'enquête soient fournies en temps opportun, afin que les décisions de politique générale puissent être fondées sur des données raisonnablement actuelles. Plus l'échantillon est grand, plus il faut de temps pour nettoyer, éditer et peser les données pour l'analyse.

F. Analyse de l'enquête

1. Établissement et ajustement des pondérations d'échantillonnage

71. Il faut affecter des pondérations aux échantillons pour compenser les probabilités de sélection inégale, les non-réponses et les différences connues entre l'échantillon et la population de référence. Ces pondérations doivent être utilisées pour l'estimation des caractéristiques étudiées de la population ainsi que pour celle des erreurs types des estimations de l'enquête.

72. La pondération de base d'une unité d'échantillon peut être assimilée au nombre d'unités de population représentées par l'unité de l'échantillon aux fins de l'estimation. Par exemple, si le taux d'échantillonnage d'une strate particulière est de 1 sur 10, la pondération de base de toute unité d'échantillon de cette strate est de 10, c'est-à-dire que l'unité de l'échantillon représente 10 unités de la population, y compris l'unité elle-même.

73. L'établissement des pondérations d'échantillonnage commence généralement par l'établissement des pondérations de base des unités de l'échantillon, pour corriger les inégalités de leurs probabilités de sélection. Dans le cas de modèles à phases multiples, la pondération de base doit tenir compte de la probabilité de sélection à chaque phase. Les pondérations de base des unités de l'échantillon sont alors ajustées pour compenser la non-réponse et la non-ouverture et pour que les estimations de l'échantillon pondéré soient conformes aux totaux connus de la population.

74. Lorsque les pondérations finales ajustées de toutes les unités de l'échantillon sont les mêmes, l'échantillon est déclaré autopondéré. Dans la pratique, les échantillons ne sont pas autopondérés et ce, pour plusieurs raisons. Premièrement, les unités d'échantillonnage sélectionnées ont la même probabilité de sélection. En fait, bien que les UPE soient souvent choisies avec une probabilité proportionnelle à leur taille, et que les ménages soient sélectionnés à un taux approprié au sein des UPE pour produire un modèle autopondéré, ces efforts peuvent être rendus vains par la sélection d'une personne à interroger au sein de chaque ménage choisi. Deuxièmement, l'échantillon choisi présente souvent des insuffisances, y compris un certain taux de non-réponse et de non-ouverture en raison des problèmes posés par le cadre d'échantillonnage (voir section C). Troisièmement, le besoin d'estimations précises pour les domaines et les sous-catégories spéciales de population exige souvent un suréchantillonnage de ces domaines (voir section D).

75. Comme on l'a indiqué précédemment, il est rare que toutes les informations désirées soient obtenues auprès de toutes les unités d'échantillonnage. Par exemple, il peut arriver que certains ménages ne fournissent aucune information, et que d'autres ne fournissent que des données partielles, c'est-à-dire des données en réponse

à certaines questions de l'enquête, mais pas à toutes. Le premier de ces types de non-réponse est appelé non-réponse unitaire ou totale, et le second, non-réponse sur certains points. Si l'on note des différences systématiques entre les personnes qui répondent et celles qui ne répondent pas, toute estimation naïve fondée uniquement sur les données recueillies auprès des personnes qui répondent ne peut être que faussée. Pour réduire ce risque de distorsion, on fait souvent des ajustements dans le cadre de l'analyse pour compenser les cas de non-réponse. La méthode type appliquée pour compenser la non-réponse sur certains points est l'imputation, qui n'est pas traitée dans ce chapitre. Voir Yansaneh, Wallace et Marker (1998), et les références citées dans leur étude, pour un examen général des méthodes d'imputation et leur application à des enquêtes vastes et complexes.

76. Pour la non-réponse unitaire, il y a trois procédures fondamentales de compensation :

- L'ajustement des pondérations de base en cas de non-réponse;
- La sélection d'un échantillon initiale plus grand qu'il n'est nécessaire, pour permettre une éventuelle réduction de la taille de l'échantillon en raison de la non-réponse;
- La substitution, qui consiste à remplacer le ménage qui ne répond pas par un autre qui n'a pas été sélectionné et qui est semblable au premier pour ce qui est des caractéristiques étudiées.

77. Il est recommandé d'utiliser une certaine forme de compensation pour la non-réponse unitaire dans les enquêtes sur les ménages, soit d'ajuster les pondérations de base des ménages qui répondent, soit par substitution. L'avantage de la substitution est qu'elle permet de conserver un certain contrôle sur le nombre de ménages participants. Toutefois, la substitution libère l'enquêteur de tout effort pour tenter d'obtenir des données auprès des ménages initialement choisis. De plus, le remplacement des ménages qui ne répondent pas prend du temps, et peut être source d'erreurs. Par exemple, on peut être tenté de substituer un ménage choisi par facilité plutôt qu'un ménage spécifiquement désigné pour servir de remplacement à un ménage qui ne répond pas. La procédure d'ajustement des pondérations d'échantillonnage pour tenir compte des cas de non-réponse est couramment utilisée dans le monde entier pour les grandes enquêtes. Par essence, l'ajustement transfère les pondérations de base de toutes les unités choisies qui ne répondent pas à des unités qui répondent. Le chapitre VIII présente une analyse plus détaillée de la non-réponse et de la non-couverture dans les enquêtes sur les ménages, et des moyens pratiques de les compenser (voir également les références citées dans ce chapitre). Le chapitre XI et les études de cas de la deuxième partie (chapitres XXII, XXIII et XXV) fournissent également des détails pour certaines enquêtes données.

78. D'autres ajustements peuvent être apportés aux pondérations en tant que de besoin. Par exemple, si l'on dispose de totaux de contrôle, on peut utiliser des ajustements post-stratification pour rendre les distributions d'échantillonnages pondérés pour certaines variables conformes aux distributions connues de la population. Voir Lehtonen et Pahkinen (1995) pour certains exemples pratiques de la façon d'analyser les données d'une enquête avec post-stratification.

2. Analyse des données d'enquêtes sur les ménages

79. Plusieurs conditions sont nécessaires à une bonne analyse des données d'une enquête sur les ménages. Premièrement, la base de données correspondante doit contenir des informations sur le processus de sélection de l'échantillon. En particulier,

elle doit comprendre des étiquettes appropriées pour désigner les strates de l'échantillon, les unités primaires d'échantillonnage, les unités secondaires d'échantillonnage, etc. Deuxièmement, des pondérations d'échantillonnage doivent être indiquées pour chaque unité du fichier de données, qui reflètent la probabilité de sélection de chaque unité d'échantillonnage et compensent la non-réponse à l'enquête et d'autres insuffisances de l'échantillon. Troisièmement, il doit y avoir suffisamment de documentation technique sur la conception de l'échantillon de l'enquête qui a produit les données. Quatrièmement, les fichiers de données doivent être d'un format et présenter une structure appropriés, et fournir les informations nécessaires sur les liens entre les unités d'échantillonnage aux divers stades de la sélection de l'échantillon. Enfin, il faut disposer du logiciel approprié, et posséder les compétences nécessaires à son utilisation.

80. Il faut un logiciel spécial pour établir des estimations des erreurs types des estimations de l'enquête qui reflètent les complexités de la conception de l'échantillon effectivement utilisé. Ces complexités sont notamment la stratification, la mise en grappes et les probabilités inégales d'échantillonnage (pondération). Les logiciels de statistiques types ne peuvent généralement pas être utilisés pour l'estimation de l'erreur type sur des modèles d'échantillon complexes, car ils supposent presque toujours que les données ont été obtenues par échantillonnage aléatoire simple. En général, l'utilisation de logiciels de statistiques types tend à sous-estimer les véritables erreurs types des estimations de l'enquête. Il existe aujourd'hui plusieurs logiciels pour l'analyse des données d'enquête obtenues à l'aide d'échantillons complexes. Certains de ces logiciels sont analysés de façon détaillée et comparés au chapitre XXI.

G. Conclusion

81. Nous concluons en soulignant quelques questions thématiques liées à la conception d'enquêtes sur les ménages dans les pays en développement et les pays en transition, à savoir :

- a) Le caractère polyvalent de la plupart des enquêtes sur les ménages : on note un regain d'intérêt, dans les pays en développement et les pays en transition, pour la mise en place de programmes permanents d'enquêtes polyvalentes, à thèmes multiples et à cycles multiples, de préférence aux enquêtes ponctuelles, spéciales. Dès le départ, le responsable de la conception de l'enquête doit reconnaître le caractère polyvalent de celle-ci et les exigences concurrentes dont ne manqueront pas de faire l'objet les données qu'elle engendrera. Ces exigences concurrentes imposent généralement des contraintes concernant l'échantillon qu'il est souvent très difficile de satisfaire. Autrement dit, le travail de l'auteur de l'enquête doit comprendre de longs entretiens avec les donateurs, les décideurs, les personnes qui produisent les données à l'office national de statistique et les usagers des données des divers ministères d'exécution du pays. Ces entretiens préliminaires ont pour objet de tenter d'harmoniser et de rationaliser les exigences concurrentes auxquelles doit répondre la conception de l'enquête, avant que l'échantillon soit finalisé;
- b) La détermination de la taille appropriée de l'échantillon : l'une des grandes questions qui se pose dès le départ a trait au choix de la taille appropriée de l'échantillon nécessaire à l'enquête. On note une demande croissante d'estimations précises des caractéristiques à étudier non seulement aux niveaux national et régional, mais aussi au niveau des provinces, voire à celui de subdivisions plus petites encore. Cela aboutit invariablement à la demande d'échantillons de grande ampleur. L'intérêt porté à assurer la fiabilité des

estimations de l'enquête en réduisant les erreurs d'échantillonnage grâce à des échantillons de grande ampleur l'emporte de beaucoup sur la nécessité tout aussi importante d'assurer la qualité des données en réduisant les erreurs non liées à l'échantillonnage. Il est recommandé à l'auteur de l'enquête d'effectuer une analyse coûts-avantages de divers choix de taille d'échantillon et de formule d'allocation. Cette analyse coûts-avantages devrait comprendre notamment un examen des erreurs non liées à l'échantillonnage dans les enquêtes et de leur impact sur la qualité d'ensemble des données de l'enquête. Les demandes d'échantillons de grande ampleur ne devraient être examinées qu'à la lumière de leurs coûts et avantages. Comme il est indiqué à la section D, il faut se rappeler, lors de l'établissement de l'échantillon, qu'il convient d'accorder la priorité aux domaines à étudier;

- c) La documentation sur la conception et l'exécution de l'enquête : dans bien des cas, la documentation sur la conception et l'exécution de l'enquête fait défaut ou est insuffisante. Pour qu'un ensemble de données soit utile aux analystes et autres utilisateurs, il est absolument essentiel que chaque aspect du processus de conception qui doit aboutir à la production de données soit documenté, notamment la sélection de l'échantillon, la collecte des données, la préparation des fichiers de données, l'établissement des pondérations d'échantillonnage ainsi que tous ajustements destinés à compenser les imperfections de l'échantillon et, si possible, les spécifications pour l'estimation des erreurs types. Nulle analyse appropriée des données ne peut être effectuée sans une telle documentation. La documentation sur l'enquête est également indispensable pour l'établissement de liens avec d'autres sources de données et pour divers types de contrôles et analyses supplémentaires;
- d) L'évaluation de la conception de l'enquête : l'un des aspects très importants du processus de conception de l'enquête a trait à la nécessité d'effectuer des analyses afin d'évaluer l'efficacité de la conception, après l'exécution. Il convient d'affecter des ressources à cet important exercice dans le cadre du processus d'établissement du budget, au stade de la planification. L'évaluation de la conception d'une enquête peut aider à améliorer la conception de l'échantillon pour les enquêtes à venir. Cette évaluation peut fournir des informations utiles sur le point de savoir si l'on a ou non tiré profit d'une allocation disproportionnée, et sur l'ampleur de la discordance éventuelle entre les mesures actuelles de la taille de l'échantillon et celles obtenues lors de sa sélection. Ces informations peuvent servir par la suite à une conception plus efficace des enquêtes.

REMERCIEMENTS

L'auteur remercie de leurs observations constructives les divers réviseurs et éditeurs, et plus particulièrement M. Graham Kalton pour ses nombreuses suggestions qui ont permis d'apporter des améliorations considérables aux versions initiales de ce chapitre. Les opinions exprimées ici sont celles de l'auteur et ne sont pas nécessairement conformes aux politiques des Nations Unies.

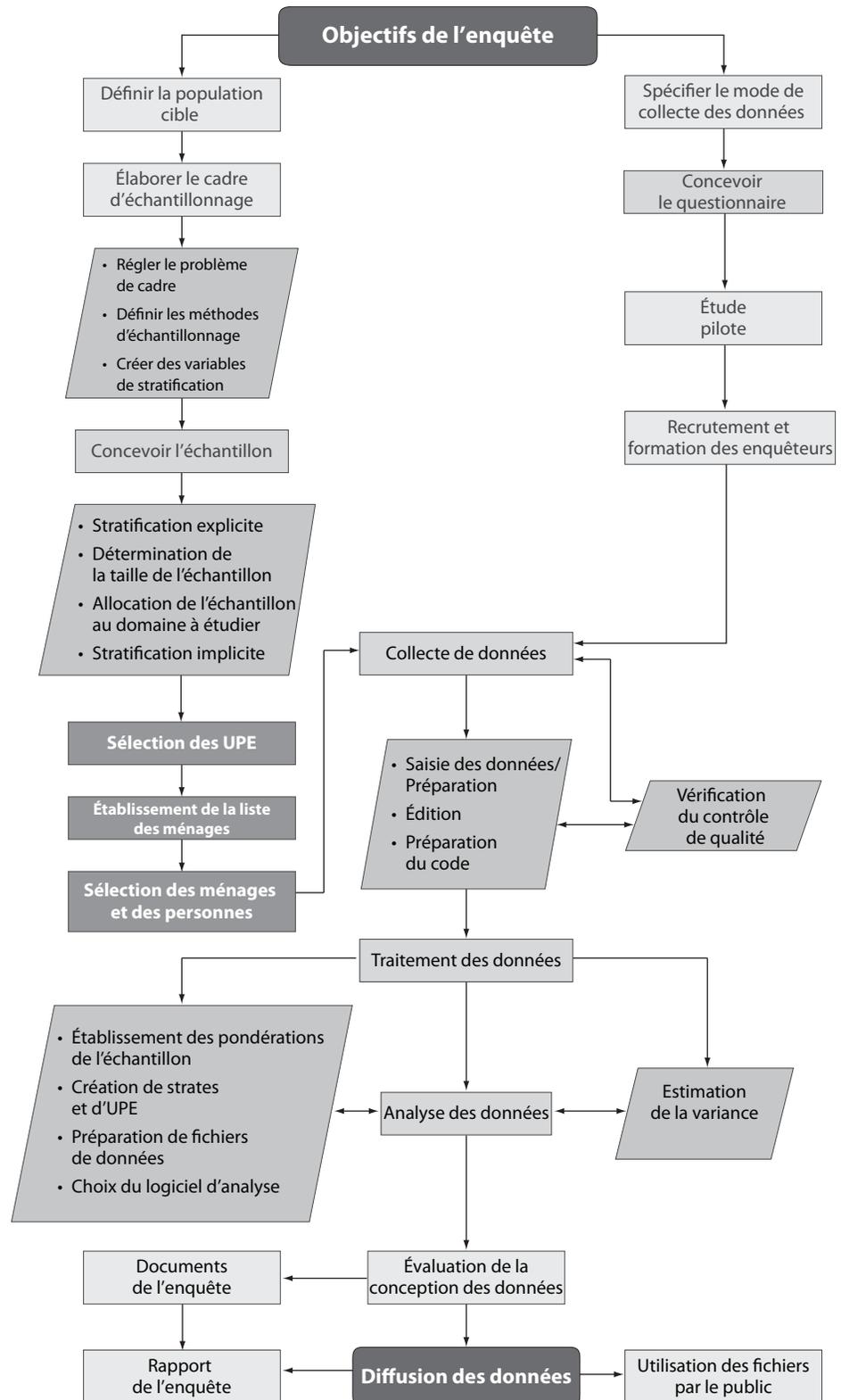
RÉFÉRENCES

- Cochran, W.G. (1977). *Sampling Techniques*, 3^e éd., New York, Wiley.
- Glewwe, P. et I. Yansaneh (2000). *The Development of Future Household Surveys in Viet Nam*. Rapport de Mission à l'Office général de statistique, Viet Nam.

- Kalton, G. (1983). *Introduction to Survey Sampling*. Quantitative Applications in the Social Sciences Series, Sage University Paper, No. 35. Beverly Hills, Californie, Sage Publications.
- Kish, L. (1965). *Survey Sampling*. New York, Wiley.
- _____ (1976). Optima and proxima in linear sample designs. *Journal of the Royal Statistical Society, Série A*, vol. 139, pp. 80-95.
- _____ (1988). Multi-purpose sample design. *Survey Methodology*, vol. 14, pp. 19-32.
- _____ (1995). Methods for design effects. *Journal of Official Statistics*, vol. 11, pp. 55-77.
- Lehtonen, R., et E. J. Pahkinen (1995). *Practical Methods for Design and Analysis of Complex Surveys*. New York, Wiley.
- Lohr, Sharon (1999). *Sampling: Design and Analysis*. Pacific Grove, Californie, Duxbury Press.
- Yansaneh, I. S. (2000). *Sample Design for the 2000 Turkmenistan Mini-census Survey*. Rapport de mission à l'Institut national de la statistique et des prévisions, Turkménistan.
- _____ (à paraître). Construction and use of sample weights. *Handbook on Household Surveys*. New York, DESA/UNSD. En préparation.
- _____, L. Wallace et D.A. Marker (1998). *Imputation methods for large complex datasets: an application to the NEHIS*. In *Proceedings of the Survey Research Methods Section, American Statistical Association*. Alexandria, Virginie, American Statistical Association. pp. 314-319.

Annexe

Diagramme du processus d'enquête



Chapitre III

Aperçu de la conception d'un questionnaire pour les enquêtes sur les ménages dans les pays en développement

PAUL GLEWWE

Département d'économie appliquée
Université du Minnesota
Saint Paul, Minnesota, États-Unis d'Amérique

RÉSUMÉ

Le présent chapitre contient un examen des questions fondamentales concernant la conception de questionnaires pour les enquêtes sur les ménages dans les pays en développement. Il commence par la première étape de cette conception, qui est la définition des objectifs de l'enquête, puis modifie ces objectifs en fonction des contraintes qui se font jour. Après un examen général de ces questions, il donne des conseils sur de nombreux aspects de la conception de tels questionnaires. Il offre également des recommandations pour les essais de ces questionnaires sur le terrain et sur leur finalisation.

Termes clés : conception du questionnaire, objectifs de l'enquête, contraintes, essai pilote, essai de terrain.

A. Introduction

1. Les enquêtes sur les ménages peuvent fournir quantité d'informations sur de nombreux aspects de la vie. Toutefois, leur utilité dépend pour beaucoup de leur qualité, c'est-à-dire aussi bien de la conception du questionnaire que de son utilisation sur le terrain. Si la conception de questionnaires pour des enquêtes et l'exécution d'enquêtes sur les ménages peuvent paraître simples à premier vue, en vérité, pour que ces enquêtes soient fructueuses, il faut leur consacrer beaucoup de temps et d'efforts.

2. Le présent chapitre donne un aperçu général du processus de conception d'un questionnaire à utiliser pour une enquête sur les ménages dans un pays en développement. Il ne s'agit que d'une introduction, car la conception d'un questionnaire est un processus très complexe qui ne peut être décrit de façon détaillée dans un chapitre de cette longueur. Nous nous bornerons donc à ébaucher les questions les plus importantes et à donner des conseils utiles à propos de chacune d'entre elles. L'étude de Grosh et Glewwe (2000) nous fournit un bon point de départ, avec des informations très détaillées sur la conception d'enquêtes sur les ménages dans les pays en développement. Bien que conçue pour un type d'enquête très précis, les études de la Banque mondiale sur la mesure des niveaux de vie, une bonne part des conseils qu'elle contient sont valables pour presque n'importe quel type d'enquête sur les ménages. De

caractère plus général, mais moins récentes, les études de Casley et Lury (1987), Nations Unies (1985), Sudman et Bradburn (1982) et Converse et Presser (1986) traitent également de la conception de ces questionnaires. On trouvera une analyse plus détaillée de la façon de concevoir une enquête sur la main-d'œuvre dans l'étude de Husmanns, Merhan et Verma (1990).

3. Tout au long de ce chapitre, on suppose que ce questionnaire sera utilisé par des enquêteurs rendant visite à des enquêtés dans leurs foyers et que l'unité d'échantillonnage est le ménage¹. Comme la plupart des enquêtes sur les ménages consistent à recueillir des informations auprès de chaque membre du ménage, elles se fondent à la fois sur des échantillons d'individus et de ménages.

4. Le reste de ce chapitre s'articule de la façon suivante : la section B est consacrée à une « vue d'ensemble », c'est-à-dire aux objectifs de l'enquête et aux difficultés auxquelles elle se heurte. La section C donne des conseils sur la structuration du questionnaire, son formatage et sur d'autres aspects de sa conception. La section D offre des recommandations sur l'ensemble du processus, depuis la constitution de l'équipe chargée de l'enquête jusqu'aux essais de terrain et à la conception finale du questionnaire. Quelques commentaires sont donnés en conclusion dans une brève section E.

B. Vue d'ensemble

5. Les questionnaires d'enquête sur les ménages varient énormément par leur contenu comme par leur longueur. La version finale de tout questionnaire est le fruit d'un processus au cours duquel des centaines, voire des milliers de décisions sont prises. On a besoin d'une « vue d'ensemble » pour s'assurer que le processus se déroule de façon ordonnée et, en fin de compte, que l'enquête atteint des objectifs qui lui ont été fixés. Pour ce faire, il faut que les concepteurs de l'enquête se mettent d'accord sur les objectifs de l'enquête et reconnaissent les difficultés auxquelles elle ne manquera pas de se heurter. La présente section explique comment établir le cadre général, en commençant par certains principes fondamentaux puis en offrant quelques conseils pratiques.

1. Objectifs de l'enquête

6. Les administrations publiques et autres organisations effectuent des enquêtes sur les ménages pour répondre aux questions qu'elles se posent au sujet de la population². Autrement dit, comme le but de l'enquête est d'obtenir des réponses à ces questions, le questionnaire doit poser les questions qui amèneront ces réponses. Étant donné les moyens et les délais limités pour répondre à l'enquête, il faut éviter toute question qui ne répond pas au but recherché. Par conséquent, la première chose à faire lors de la conception d'une enquête sur les ménages est de convenir de ses objectifs et de les définir par écrit.

7. Pour fixer les objectifs de l'enquête, les concepteurs doivent commencer par identifier un ensemble de questions auxquelles l'organisme parrainant l'enquête désire avoir des réponses. Quatre types de questions sont à considérer. Le plus simple a trait aux questions sur les caractéristiques fondamentales de la population à l'heure actuelle. Voici quelques exemples de ces questions :

- *Quel est le pourcentage de la population pauvre ?*
- *Quel est le taux de chômage ?*
- *Quel est le taux de malnutrition chez les jeunes enfants ?*

¹ Dans certaines enquêtes, l'unité d'échantillonnage est le logement, non le ménage, mais en pareils cas, une partie ou la totalité des ménages des logements de l'échantillon deviennent les « unités déclarantes » de l'enquête. Par ailleurs, certaines populations étudiées ne peuvent être couvertes par l'enquête sur les ménages. Tel est le cas des enfants des rues et des nomades. Néanmoins, la plupart des informations contenues dans ce chapitre s'appliquent à ces types de population. Pour plus de détails sur la façon d'enquêter auprès de ces populations, voir Nations Unies (1993).

² Ces questions générales auxquelles l'organisation qui effectue l'enquête aimerait avoir des réponses ne sont pas nécessairement les mêmes que les questions plus spécifiques contenues dans le questionnaire de l'enquête et posées aux membres des ménages. La présente section porte sur le premier de ces deux types de questions.

- *Quels sont les produits cultivés par les familles rurales des différentes régions du pays ?*

8. Le deuxième type de questions doit tendre à établir un lien entre les caractéristiques des ménages et les politiques et programmes du gouvernement, de manière à déterminer le taux de couverture de ces programmes. Voici un exemple de ce genre de questions :

- *Quel est le pourcentage de ménages qui participent à un programme donné, et quelle est la différence de caractéristiques entre ces ménages et ceux qui ne participent pas à ce programme ?*

9. Le troisième type de question a trait aux *changements* observés dans le temps dans les caractéristiques des ménages. Les administrations publiques et autres organisations cherchent souvent à savoir si les conditions de vie des ménages s'améliorent ou se dégradent. Il faut les données de deux enquêtes ou plus, considérablement espacées dans le temps, pour répondre à ce type de questions, et il faut que les données recherchées soient recueillies de la même manière lors de chaque enquête. Comme l'expliquent Deaton et Grosh (2000), même des façons légèrement différentes de recueillir ces données peuvent produire des résultats qui ne sont pas comparables et qui peuvent être trompeurs.

10. Le quatrième et dernier type de questions a trait aux déterminants (causes) de la situation et des caractéristiques des ménages. Il est difficile de répondre à ces questions, car il ne s'agit pas seulement de demander *qu'est-ce qui se passe* mais *pourquoi* cela se passe. Or ces questions sont souvent les plus importantes car elles cherchent à expliquer l'impact des politiques et programmes actuels, voire des hypothétiques politiques et programmes à venir, sur la situation et les caractéristiques des ménages. Les économistes et autres experts en sciences sociales ne sont pas toujours d'accord sur la façon de répondre à ces questions, et certains peuvent même parfois douter qu'il soit possible de répondre à une question particulière. Si ces questions sont importantes pour les concepteurs de l'enquête, il faut qu'elles fassent l'objet d'une planification minutieuse. Toutefois, les problèmes que pose cette planification débordent le cadre de ce chapitre (voir divers chapitres de Grosh et Glewwe (2000) pour un examen détaillé de ce qu'il faut pour répondre à ce type de question).

11. Une fois qu'un ensemble de questions auxquelles répondre est convenu, ces questions peuvent être l'expression des objectifs de l'enquête. Par exemple, la présence d'une question sur le taux de chômage actuel implique que l'un des objectifs de l'enquête est de mesurer l'incidence du chômage sur la population économiquement active. L'étape suivante est de classer ces objectifs par ordre d'importance. Si le nombre d'objectifs est élevé, il est très possible que l'enquête ne puisse recueillir toutes les informations nécessaires pour les atteindre tous, en raison d'insuffisances budgétaires, de limites de capacité et d'autres contraintes. Lorsque tel est le cas, les objectifs moins prioritaires (en regard de l'effort à accomplir pour recueillir les informations nécessaires pour y répondre) devraient être abandonnés³. Lors du choix des objectifs auxquels doit répondre l'enquête, il convient de vérifier si d'autres données déjà disponibles peuvent être utilisées pour répondre à la question liée à tel ou tel objectif. Tout objectif auquel il est possible de répondre en utilisant des données d'autres sources devrait être retiré de la liste des objectifs de la nouvelle enquête. Le processus de sélection d'un ensemble raisonnable d'objectifs est plus un art qu'une science, et les concepteurs de l'enquête doivent aussi tenir compte de certains facteurs tels que l'expérience passée concernant la collecte de données pertinentes pour un objectif et de la capacité d'ensemble de l'organisme effectuant l'enquête. Une fois que ces difficultés sont apla-

³ Au lieu d'abandonner un objectif moins important, on pourrait chercher à recueillir les données nécessaires pour y répondre en n'interrogeant qu'un sous-échantillon de ménages. Cela nécessiterait moins de ressources, mais diminuerait aussi la précision des estimations et pourrait en même temps compliquer l'exécution de l'enquête sur le terrain.

nies, cette approche devrait permettre aux concepteurs de l'enquête sur les ménages de convenir de la liste des objectifs que cherche à atteindre cette enquête.

12. Le dernier point à signaler est que certains concepteurs d'enquêtes préfèrent présenter l'ensemble de leurs questions ou de leurs objectifs sous la forme de tableaux à remplir à l'aide des réponses obtenues. Cette approche, souvent appelée « plan de tabulation », est plus efficace avec les trois premiers types de questions. Généralement, la façon dont les données recueillies lors d'une enquête sur les ménages sont utilisées pour répondre aux questions posées (atteindre les objectifs) est désignée sous le nom de « plan d'analyse des données ». Ces plans, qui peuvent être très détaillés, devraient être conçus une fois que les détails de l'enquête sur les ménages ont été convenus (ce point est examiné de façon plus détaillée à la section C).

2. Contraintes

13. Le processus de sélection des objectifs décrit ci-dessus est subordonné à un ensemble de contraintes qui limitent ce qui peut être fait. Les concepteurs d'enquêtes se heurtent à trois principales contraintes. La première et la plus évidente a trait aux ressources financières dont ils disposent. Cette contrainte limite à la fois le nombre de ménages à interroger et le temps que les enquêteurs peuvent passer avec chaque ménage (ce qui limite à son tour le nombre de questions qui peuvent être posées à un ménage donné). En général, il y a différentes combinaisons possibles entre la taille de l'échantillon (le nombre de ménages à interroger) et la quantité d'informations que l'on peut obtenir auprès de chaque ménage, et pour un budget donné, il y a des possibilités d'arbitrage entre ces deux aspects de l'enquête. En particulier, pour un volume donné de ressources financières, on ne peut accroître la taille de l'échantillon qu'en diminuant la quantité d'informations recueillies auprès de chaque ménage, et vice versa⁴. Il est clair que cela ne peut manquer de se répercuter sur le nombre d'objectifs de l'enquête et sur la précision de ces objectifs (c'est-à-dire la précision des réponses aux questions posées) : un petit échantillon permet de recueillir plus de données par ménage et, donc, de répondre aux questions posées, mais la précision de ces réponses est plus faible du fait de la taille plus réduite de l'échantillon. Ainsi, la qualité des données, c'est-à-dire la précision des informations, est également fonction des ressources disponibles. Par exemple, si l'on dispose de fonds pour laisser plus de temps à chaque enquêteur pour remplir un questionnaire d'une taille donnée, il peut utiliser ce surcroît de temps pour retourner chez les ménages afin de corriger certaines erreurs ou incohérences dans les données qu'il peut avoir détectées une fois que l'entrevue est terminée.

⁴ Pour un budget donné, la relation exacte entre les informations recueillies par ménage et le nombre de ménages interrogés est rarement simple. En particulier, il n'est pas vrai que l'on puisse, par exemple, doubler la taille de l'échantillon en réduisant de moitié celle du questionnaire, pour une durée donnée de temps de l'enquêteur. En effet, les enquêteurs ont besoin de beaucoup de temps pour trouver les ménages, se présenter, puis passer au ménage ou à la zone de recensement suivante, et l'on ne peut réduire ce temps en raccourcissant le questionnaire.

14. La deuxième contrainte à laquelle doivent faire face les concepteurs d'enquêtes a trait à la capacité de l'organisme chargé d'effectuer l'enquête. Les échantillons de grande taille ou les questionnaires très détaillés peuvent dépasser l'aptitude de cet organisme à entreprendre l'enquête au niveau de qualité désiré. Plus l'échantillon est grand, plus le nombre d'enquêteurs et de personnes chargées d'entrer les données qu'il faut engager et former est élevé, ce qui signifie que l'organisme peut avoir à réduire le minimum de qualifications acceptables de la part de ces personnes s'il veut pouvoir engager l'effectif requis. Par ailleurs, des questionnaires plus longs nécessitent plus de formation et un personnel plus compétent, et les enquêteurs, de même que les personnes chargées d'entrer les données ayant reçu la formation et présentant les qualifications requises font souvent défaut dans les pays en développement. Cette contrainte n'est pas toujours pleinement admise, de sorte que nombre d'enquêtes effectuées dans ces pays produisent de vastes ensembles de données de qualité douteuse et par conséquent d'une utilité discutable.

15. La dernière contrainte a trait à l'inclination et à l'aptitude des ménages interrogés à fournir les informations désirées. Premièrement, les ménages ne sont guère enclins à répondre aux questions, de sorte que les questionnaires très longs ont des chances de se traduire par des taux élevés de refus et/ou de données incomplètes ou inexactes. Deuxièmement, même lorsque les enquêtés sont prêts à coopérer, ils ne sont pas toujours à même de répondre à des questions complexes ou à des questions portant sur les événements qui se sont produits des mois, sinon des années auparavant. Cela a des répercussions directes sur la conception du questionnaire. Par exemple, il peut se révéler impossible d'obtenir une estimation relativement précise du revenu d'un ménage en posant un petit nombre de questions, et il peut être nécessaire de poser une longue série de questions détaillées; tel est particulièrement le cas des familles d'agriculteurs des zones rurales qui pratiquent de nombreuses cultures dont certaines servent à leur consommation et d'autres à la vente.

3. Quelques conseils pratiques

16. Les concepteurs d'enquêtes doivent revenir plusieurs fois sur les objectifs de l'enquête et les contraintes auxquelles ils se heurtent jusqu'à ce qu'ils parviennent à un ensemble d'objectifs qui soient à la fois réalisables, étant donné ces contraintes, et « optimaux » en ce sens qu'ils sont les plus importants pour l'organisme qui effectue l'enquête. Une fois que ce qui est réalisable devient clair, il peut être possible d'alléger les contraintes en mobilisant des ressources financières supplémentaires ou en offrant une formation supplémentaire aux futurs enquêteurs. L'expérience d'autres enquêtes effectuées récemment dans le même pays devrait permettre de se faire une idée de ce qui est possible et de ce qui est irréaliste. Comme on l'a déjà indiqué, trouver le juste équilibre relève davantage de l'art que de la science, et les expériences locale et internationale devraient aider à parvenir à cet équilibre.

C. Les détails

17. Une fois que l'on a une « vue d'ensemble » des objectifs de l'enquête, les concepteurs doivent entreprendre la tâche laborieuse de la conception détaillée du questionnaire, question par question. Tout d'abord, il faut un plan d'analyse des données. Ce plan doit expliquer de façon détaillée quelles sont les données nécessaires pour atteindre les objectifs fixés pour l'enquête (les réponses aux questions). Les concepteurs doivent se référer constamment à ce plan lorsqu'ils travaillent aux détails du questionnaire. Dans certains cas, le plan d'analyse des données doit être modifié à mesure que se précise la façon dont les données devraient être analysées. Toute question qui n'est pas utilisée par ce plan doit être retirée du questionnaire.

18. Ce chapitre est beaucoup trop bref pour répondre de façon détaillée à la question de savoir comment concilier la conception du questionnaire avec certains objectifs spécifiques et les plans d'analyse des données correspondants. Pour des conseils plus détaillés applicables aux différents types d'enquêtes, voir les divers chapitres consacrés à certains thèmes précis dans Grosh et Glewwe (2000). Le reste de la présente section offre quelques conseils généraux mais très utiles sur la façon de régler les détails de l'élaboration d'un questionnaire d'enquête sur les ménages.

1. L'approche par modules

19. Un questionnaire d'enquête sur les ménages se compose généralement de plusieurs parties, souvent appelées modules. Un module est constitué d'une ou de plusieurs pages de questions sur un sujet particulier, tel que le logement, l'emploi ou la

santé. Par exemple, la série d'Enquêtes démographiques et sanitaires examinées au chapitre XXII a des modules sur la contraception, les préférences en matière de fécondité et la vaccination des enfants. D'une façon générale, dans presque tout questionnaire d'enquête sur les ménages ayant plusieurs questions sur un sujet donné, tel que l'éducation de chaque membre de la famille, il est plus commode de poser ces questions ensemble sur une ou plusieurs pages du questionnaire et de désigner cette page ou ces pages comme le module sur ce sujet; par exemple, les questions sur l'éducation deviendraient le « module éducation ». Ainsi, le questionnaire dans son ensemble peut être perçu comme une collection de modules, dont le nombre peut être de trois seulement ou de 15 ou 20, selon le nombre de thèmes couverts par le questionnaire. Chaque module comporte plusieurs questions, parfois 5 ou 6 seulement, et parfois 50, voire plus de 100⁵. Les très gros modules, tels que ceux qui comptent plus de 50 questions, doivent être eux-mêmes subdivisés en sous-modules portant sur un sujet particulier : emploi principal, emploi secondaire, antécédents professionnels. Quoi qu'il en soit, le nombre total de questions d'un questionnaire doit être maintenu au minimum nécessaire pour produire les informations désirées.

⁵ Un module de plus de 100 questions peut nécessiter une entrevue trop longue. Voir section D pour un examen plus approfondi de la longueur totale du questionnaire.

20. L'approche par modules est pratique, car elle permet de fractionner la conception du questionnaire en deux étapes : la première étant consacrée à décider des modules nécessaires, c'est-à-dire des sujets à traiter, et de l'ordre dans lequel les faire figurer sur le questionnaire; et la seconde consistant à concevoir chaque module, question par question. Pendant ces deux étapes, il faut se référer constamment aux objectifs de l'enquête et au plan d'analyse des données.

21. Le choix des modules et les détails de chaque module sont très variables, selon les objectifs de l'enquête et les contraintes auxquelles elle se heurte. Cependant, on peut donner quelques conseils d'ordre général applicables à presque toutes les enquêtes. Par exemple, presque toutes les enquêtes sur les ménages contiennent des questions sur le nombre de personnes composant la famille, et certaines questions très élémentaires sur leur âge, leur sexe et leur lien avec le chef de famille. Ces questions peuvent être regroupées sur un bref module d'une page sur la « composition de la famille ». Ce module devrait être l'un des tout premiers, et le plus souvent le premier, du questionnaire. Beaucoup de questionnaires d'enquête sur les ménages contiennent ensuite des questions à poser à chaque membre de la famille sur des sujets tels que l'éducation, l'emploi, la santé et la migration. Chacun de ces sujets sur lesquels il est posé cinq questions ou plus devrait probablement faire l'objet d'un module spécial sur ce sujet. S'il ne fait l'objet que d'une, deux ou trois questions, il pourrait être plus pratique d'inclure celles-ci dans le module sur la composition de la famille, ou peut-être dans un autre module contenant des questions posées individuellement à chaque membre de la famille.

22. Presque tous les modules d'un questionnaire d'enquête sur les ménages peuvent être classés en deux grandes catégories : ceux qui contiennent des questions posées à chaque membre de la famille, et ceux qui contiennent des questions d'ordre général sur la famille. En ce qui concerne la première catégorie, il convient de préciser que les questions posées à chaque membre de la famille n'ont pas besoin d'être les mêmes pour tous; beaucoup de questionnaires d'enquête sur les ménages contiennent des questions qui ne s'appliquent qu'à certains membres de la famille, telles que le nombre d'enfants de moins de cinq ans ou de femmes en âge de procréer. Parmi les exemples de la seconde catégorie figurent les questions sur les caractéristiques du logement où vit la famille et les questions sur les dépenses familiales consacrées à l'alimentation et les autres. Bien entendu, la longueur de chacun de ces modules et les types de questions qu'ils contiennent sont fonction des objectifs de l'enquête.

23. Enfin, on peut faire quelques remarques d'ordre général sur l'ordre dans lequel doivent figurer les modules dans le questionnaire. Premièrement, cet ordre doit être celui dans lequel les questions seront posées lors de l'entrevue, afin que l'enquêteur puisse remplir le questionnaire en commençant à la première page et en poursuivant page par page jusqu'à la fin du questionnaire. Il peut y avoir des exceptions à cette règle, mais en général, il est « naturel » de procéder ainsi.

24. Deuxièmement, les premiers modules du questionnaire devraient contenir des questions auxquelles il soit relativement aisé de répondre et qui traitent de sujets qui ne soient pas délicats. La suggestion précédente d'utiliser la composition de la famille comme premier module est conforme à cette recommandation, car les informations de base sur les membres de la famille ne sont généralement pas un sujet délicat. Le fait de commencer l'entrevue par des questions simples sur des sujets non délicats permet à l'enquêteur de mettre les membres de la famille à l'aise et d'établir le contact avec eux. Cela implique que les modules les plus délicats devraient être relégués à la fin du questionnaire. Cela laisse à l'enquêteur le plus de temps possible pour gagner la confiance des membres de la famille, et augmente la probabilité que ceux-ci répondent pleinement et correctement aux questions délicates. De plus, si des questions délicates poussent les membres de la famille à mettre fin à l'entrevue, l'enquêteur aura au moins obtenu les réponses à toutes les questions non délicates.

25. Le troisième principe consiste à grouper les modules dont les questions ont des chances de recueillir des réponses du même membre de la famille. Par exemple, les questions sur les dépenses consacrées à l'alimentation et les autres devraient être groupées, car il y a des chances qu'un membre de la famille soit mieux à même que les autres d'y répondre. Cela permet à ce membre de la famille de répondre à toutes les questions de ces modules auxquelles il ou elle peut répondre, puis de se retirer et de laisser à d'autres membres de la famille le soin de répondre aux questions des autres modules. Le but est d'utiliser efficacement le temps des membres de la famille, ce qu'ils apprécient et qui les incite à coopérer davantage. Cela devrait également faire gagner du temps à l'enquêteur, car chaque déclarant n'a besoin d'être interrogé qu'une fois pour apporter sa contribution à l'enquête.

2. Formatage et homogénéité

26. Une fois que les modules ont été sélectionnés et que leur ordre a été déterminé, il faut passer au stade délicat et laborieux du choix des questions et de leur rédaction. Avant de commencer, il est prudent de regarder les précédentes enquêtes sur les ménages effectuées dans le pays ou, éventuellement, dans un pays voisin, sur le même sujet. Le plus souvent, si la qualité des questions et leur rédaction sont fonction de la nature et de l'objet de la nouvelle enquête, on peut néanmoins donner quelques conseils d'ordre général valables pour presque toutes les enquêtes sur les ménages.

27. La première recommandation est que, dans la plupart des cas, les questions soient rédigées de manière que l'enquêteur puissent les poser simplement en les lisant sur le questionnaire. Ainsi, les questions peuvent être posées dans les mêmes termes à tous les ménages. Dans le cas contraire, le questionnaire serait rédigé en termes succincts, obligeant l'enquêteur à les formuler dans ses propres termes. Cela est à éviter en raison des risques d'erreurs d'interprétation. Par exemple, supposons qu'un module sur l'emploi contienne une question demandant simplement quelle est la « principale occupation ». Cela peut prêter à confusion. S'agit-il de la principale occupation du jour ou de la semaine de l'entrevue, ou des 12 derniers mois ? Pour les personnes qui exercent deux métiers, le principal est-il celui qui lui rapporte le plus gros revenu ou celui auquel il consacre le plus d'heures ou de jours de travail ? On peut éviter cette

confusion si la question est rédigée de façon plus précise, comme dans l'exemple suivant : « Au cours des sept derniers jours, quel type de travail avez-vous fait ? Si vous avez travaillé à plusieurs tâches, à laquelle avez-vous consacré le plus d'heures au cours des sept derniers jours ? » La figure III.1 offre un exemple de la page de questionnaire destinée à recueillir des informations sur le logement (on remarquera que toutes les questions sont posées sous la forme de phrases complètes). L'avantage de cette façon de poser les questions a été clairement démontré dans une étude expérimentale de Scott *et al.* (1988) : les questions qui n'avaient pas été posées sous la forme de questions détaillées ont été la source de 7 à 20 fois plus d'erreurs que celles qui avaient été rédigées en phrases précises.

28. La deuxième recommandation est étroitement liée à la première : le questionnaire doit donner des définitions précises de tous les principaux concepts utilisés, essentiellement pour permettre à l'enquêteur de se référer à ces définitions durant l'entrevue lorsque se présentent des cas inhabituels. De plus, il doit contenir des commentaires explicatifs à l'intention de l'enquêteur; on trouvera des exemples de tels commentaires à propos de la question 10 de la figure III.1. Des instructions et explications plus détaillées devraient faire l'objet d'un manuel à l'usage de l'enquêteur. Ce type de manuel est traité au chapitre IV.

29. La troisième recommandation est de garder les questions aussi courtes et simples que possible, et d'utiliser des termes courants. Par ailleurs, il faut vérifier soigneusement chaque question pour s'assurer qu'elle n'est pas « tendancieuse » ou qu'elle ne risque pas d'inciter le déclarant à donner des réponses partiales. Si la question est compliquée, il est recommandé de la découper en deux questions distinctes ou plus. Voici un exemple, à titre d'illustration. Supposons que l'on cherche à savoir si une personne était employée ou « travailleur indépendant » (ou les deux) au cours des sept derniers jours. Essayer d'obtenir ces informations à l'aide d'une seule question en utilisant un jargon plus ou moins technique pourrait donner le libellé suivant :

- *Au cours des sept derniers jours, avez-vous été employé salarié ou percevant une autre forme de rémunération ou travailleur indépendant d'une entreprise familiale ? Vous êtes-vous livré simultanément à ces deux types d'activités, ou à aucune d'entre elles ?*

Au lieu d'une telle question, il conviendrait de poser deux questions séparées en utilisant des termes moins techniques :

1. *Au cours des sept derniers jours, avez-vous travaillé pour de l'argent pour quelqu'un qui n'est pas membre de votre famille ?*
2. *Au cours des sept derniers jours, avez-vous travaillé à votre compte, par exemple comme agriculteur ou comme marchand de biens ou de services ?*

Les questions 8, 9 et 10 de la figure III.1 offrent une autre illustration de ce point. Les rédacteurs du questionnaire peuvent être tentés « de raccourcir » celui-ci en combinant ces questions en une seule, de la façon suivante :

- *De quel type de titre ou document légal éventuel disposez-vous, attestant que vous êtes propriétaire de ce logement, et quel membre de la famille détient effectivement ce titre ?*

Ce type de question plus longue pourrait troubler de nombreux déclarants, et si tel est le cas, il pourrait falloir plus de temps pour expliquer la question que pour en poser trois séparément.

30. Quatrièmement, le questionnaire devrait être conçu de manière que les réponses à presque toutes les questions soient précodées. Ces questions sont souvent appelées « questions fermées » par les concepteurs de l'enquête. Par exemple, les ré-

Figure III.1
Illustration du formatage d'un questionnaire

1. Ce logement appartient-il à un membre de votre famille ?

Oui 1

Non 2 (»12)

2. Comment votre famille a-t-elle obtenu ce logement ?

CONVERSION 1

ACHAT À UN PARTICULIER 2

NOUVELLE CONSTRUCTION 3

COOPÉRATION 4

ÉCHANGE 5 (»7)

HÉRITAGE 6 (»7)

AUTRE 7 (»7)

3. Comment avez-vous payé ce logement ?

4. Payez-vous votre logement par tranches ?

Oui 1

Non 2 (»7)

5. Quel est le montant des tranches ?

MONTANT (UNITÉS MONÉTAIRES)

PÉRIODICITÉ

6. En quelle année comptez-vous effectuer votre dernier paiement ?

ANNÉE

7. Possédez-vous un titre de propriété sur le terrain ou un autre document attestant que vous en êtes propriétaire ?

Oui 1

Non 2

8. Possédez-vous un titre de propriété sur le logement ou un autre document attestant que vous en êtes propriétaire ?

Oui 1

Non 2

9. De quel type de titre s'agit-il ?

TITRE PLEINEMENT LÉGAL, ENREGISTRÉ 1

TITRE LÉGAL, NON ENREGISTRÉ 2

RÉCIPISSE D'ACHAT 3

AUTRE 4

10. Qui détient ce titre ou document dans la famille ?

ÉCRIRE LE CODE D'IDENTIFICATION SI CETTE PERSONNE FIGURE SUR LA LISTE

PREMIER CODE

SECOND CODE

11. Pourriez-vous vendre ce logement si vous le désiriez ?

Oui 1

Non 2 (»14, PAGE SUIVANTE)

12. Si vous vendiez ce logement aujourd'hui, quel montant en tireriez-vous ?

MONTANT (UNITÉS MONÉTAIRES)

13. Veuillez donner une estimation du montant du loyer que vous pourriez percevoir si vous décidiez de louer ce logement ?

MONTANT (UNITÉS MONÉTAIRES)

PÉRIODICITÉ

» QUESTION 28, PAGE SUIVANTE								
UNITÉS DE TEMPS :	JOURNÉE	3	MOIS	6	ANNÉE ..	9		
	SEMAINE	4	TRIMESTRE ...	7				
	QUINZAINE ...	5	SEMESTRE	8				

⁶ Une autre solution serait de permettre à l'enquêteur d'inscrire un « X » ou un signe dans une case à côté d'une réponse précodée.

ponses aux questions auxquelles il suffit de répondre par *oui* ou par *non* peuvent être enregistrées sur le questionnaire comme « 1 » pour *oui* et « 2 » pour *non*. Cela est plus facile pour l'enquêteur qui n'a à écrire qu'un seul chiffre au lieu d'un mot ou d'une phrase⁶. Mais surtout, cela permet d'éliminer la phase de « codage », au cours de laquelle des questionnaires contenant des réponses manuscrites (souvent illisibles) d'un ou de plusieurs mots sont remis à un agent « codeur » qui doit alors inscrire un code numérique pour chacune de ces réponses. Cette opération supplémentaire est une nouvelle source d'erreurs, et peut être évitée dans la plupart des cas. (En revanche, le codage de classifications plus complexes, telles que la profession et le secteur d'activité, nécessite des qualifications et un temps dont les agents de terrain disposent rarement, et il est recommandé qu'il soit confié à des codeurs qualifiés opérant à partir des descriptions écrites données par les enquêteurs.) À la figure III.1, toutes les réponses possibles aux questions sont précodées, et tous les codes sont indiqués sur la même page que la question (le plus souvent immédiatement après la question).

31. La cinquième recommandation est liée à la troisième. Le codage des réponses doit être homogène pour toutes les questions. Par exemple, presque toutes les enquêtes sur les ménages contiennent des questions auxquelles il faut répondre par *oui* ou par *non*. Le code numérique utilisé pour toutes ces questions doit toujours être le même, par exemple « 1 » pour *oui* et « 2 » pour *non*. Une fois cette règle (ou une autre) établie, elle doit être utilisée pour toutes les réponses par *oui* ou par *non* aux questions du questionnaire. Cette règle peut être appliquée également à d'autres types de réponses. Beaucoup de questionnaires ont des questions auxquelles les réponses doivent s'exprimer en unités de temps ou de distance, comme, par exemple, « Quand avez-vous vu un médecin pour la dernière fois ? » ou « À quelle distance votre maison se trouve-t-elle de la route la plus proche ? » Les unités de temps pourraient être codées de la façon suivante : 1 pour les minutes, 2 pour les heures, 3 pour les jours, 4 pour les semaines, et ainsi de suite. Ainsi, une réponse de « 10 jours » serait enregistrée au moyen de deux chiffres : « 10 » et « 3 » où 3 serait l'unité codée de temps. De la même façon, pour les distances, 1 pourrait indiquer les mètres et 2 les kilomètres. La formule de codage peut être différente d'une enquête à une autre; l'important est que, dans la mesure du possible, toutes les questions nécessitant un code de ce type utilisent le même codage⁷. La figure III.1 illustre aussi cette recommandation. On remarquera que les codes d'unités de temps donnés au bas de la page s'appliquent à deux questions de cette page, à savoir les questions 5 et 13.

⁷ S'il importe peu que les chiffres utilisés comme codes pour certains concepts simples, tels que les unités de temps et de distance, diffèrent d'une enquête à une autre dans le même pays, il est préférable d'utiliser le même codage pour les concepts plus complexes, tels que les types de professions ou les types de maladies, afin que les résultats d'enquêtes effectuées à différentes périodes puissent être comparés.

32. Ces réflexions sur les formules de codage amènent à se poser la question de savoir si l'enquêteur doit indiquer aux déclarants les réponses possibles aux questions ou se contenter de lire les questions en omettant les codes. En général, la seconde méthode est préférable. Les déclarants peuvent donner l'une des réponses simplement parce qu'ils ont entendu d'abord cette réponse, même si une autre réponse est plus conforme à la réalité. De plus, s'il y a un choix entre un grand nombre de réponses possibles, les déclarants peuvent commettre des erreurs et choisir une mauvaise réponse.

33. La sixième recommandation est que le questionnaire contienne des « codes à omettre » indiquant quelles questions ne pas poser à la famille, sur la base de ses réponses aux questions précédentes. Par exemple, un questionnaire peut contenir la question suivante : « Avez-vous cherché du travail au cours des sept derniers jours ? » Si la réponse est *oui*, les questions suivantes peuvent avoir trait aux méthodes employées, mais si la réponse est *non*, de telles questions deviennent sans objet. De très brèves instructions, telles que « SI LA RÉPONSE EST NON, PASSER À LA QUESTION 6 » devraient figurer juste à côté de la première question, afin que l'enquêteur ne pose pas les questions suivantes qui deviennent sans objet. On pourrait adopter certaines conventions pour donner ces instructions plus succinctement, par exemple :

« SI NON → Q.6 ». À la figure III.1, les instructions régies par ces conventions sont très brèves, elles sont données par des chiffres entre parenthèses après les réponses codées pertinentes. Par exemple, la marque « (»12) » après le code NON à la question 1 indique que si la réponse à cette question est *non*, l'enquêteur doit passer à la question 12.

34. Enfin, en ce qui concerne le formatage, il convient de préciser que les questions doivent être posées d'une manière qui permette au déclarant d'y répondre dans ses propres termes. Voici un exemple qui permettra de mieux comprendre ce point : dans une enquête sur le logement, il peut y avoir une question sur le loyer payé par le ménage. Selon le contrat de location, certains déclarants peuvent payer leur loyer à la semaine, d'autres au mois et d'autres encore à l'année. Il s'agit donc ici le permettre au répondant de choisir l'unité de temps, de sorte que la question devrait être « Combien payez-vous de loyer pour votre logement ? » et non pas « Combien payez-vous par mois de loyer pour votre logement ? » Le problème avec cette dernière question est qu'elle force le déclarant à répondre en chiffres mensuels. Or, celui-ci peut très bien savoir qu'il paie \$50 par semaine, mais faire une erreur en multipliant \$50 par 4,3 (\$217 par mois). Il vaut mieux rédiger le questionnaire de manière que l'enquêteur puisse inscrire des codes numériques pour les différentes unités de temps, comme l'illustre la question 5 de la figure III.1, de manière que \$50 par semaine, par exemple, puisse être enregistré comme 50 dans un espace plus 4 (le code numérique pour la semaine) dans l'espace voisin. Lorsque les données seront analysées, l'enquêteur, qui risque beaucoup moins de commettre une erreur que le déclarant, pourra facilement convertir les montants en une unité commune, telle que le loyer annuel.

3. Autres conseils sur certains aspects de la conception du questionnaire

35. Pour terminer, voici quelques conseils d'ordre plus général au sujet de la conception du questionnaire. Premièrement, pour les questions très importantes, telles que le nombre de personnes composant la famille ou les sources de revenu du ménage, il peut être opportun de poser une question « test » qui permette au déclarant de se souvenir d'un détail qu'il ou elle a pu oublier. Par exemple, après avoir obtenu la liste des membres de la famille, l'enquêteur peut poser la question suivante :

- *D'après les informations que vous m'avez données, votre famille compte six membres. Est-ce exact ou quelqu'un d'autre fait-il également partie de la famille, par exemple quelqu'un qui peut être absent pour quelques jours ou quelques semaines ?*

36. Deuxièmement, le questionnaire devrait être conçu de manière que chaque famille et chaque membre de la famille ait un numéro de code qui permette de l'identifier tout au long du questionnaire. Cela aiderait les analystes des données à associer les informations correspondant à une même famille ou à une même personne. Dans la plupart des cas, il devrait y avoir un questionnaire par famille; dans les cas exceptionnels où l'on utilise deux questionnaires ou plus, il faut veiller d'autant plus à ce que le même code de famille figure sur chacun des questionnaires remplis pour une même famille.

D. Le processus

37. Jusqu'ici, nous avons donné des conseils sur la façon de concevoir les questionnaires d'enquête sur les ménages, mais nous n'avons presque rien dit de leurs concepteurs ni de la façon dont ils peuvent vérifier le questionnaire qu'ils ont conçu.

La présente section contient des recommandations concernant la rédaction, la vérification et la finalisation du questionnaire.

1. Constitution d'une équipe

38. Les enquêtes sur les ménages nécessitent presque toujours un très grand nombre de décisions et d'initiatives, qui se révèlent généralement plus compliquées que prévu. Cela veut dire qu'une seule personne ou même un petit groupe de personnes peuvent ne pas avoir assez de temps ou de compétences pour concevoir un bon questionnaire. C'est pourquoi il convient de former une équipe « d'experts » dès le début du processus, afin qu'aucun aspect de l'enquête ne soit négligé. Cette équipe devrait être composée de représentants de plusieurs groupes clés.

39. Il peut être très important de compter un ou plusieurs membres du groupe de décideurs dans l'équipe, c'est-à-dire une ou plusieurs personnes représentant les intérêts du groupe ou des groupes qui envisagent d'utiliser les informations recueillies dans le cadre de l'enquête pour décider des orientations à suivre. Bien que ces personnes ne soient pas des experts techniques, elles sont nécessaires pour informer les autres membres de l'équipe des objectifs ultimes de l'enquête (ou pour les leur rappeler). La présence de ce groupe dans l'équipe augmente sensiblement la communication entre les producteurs des données et leurs usagers.

40. Le deuxième groupe clé, composé de chercheurs et d'analystes de données, utilise les informations que contiennent ces données pour répondre aux questions intéressant les décideurs. Ces personnes ont pour rôle d'élaborer le plan d'analyse des données qui permettra d'assurer que les données recueillies répondent à ces questions. Dans certains cas, il est facile de répondre à ces questions, mais dans d'autres, cela peut être très compliqué.

41. Le dernier mais non le moindre est le groupe de personnes chargées de recueillir les données, c'est-à-dire les enquêteurs, les superviseurs et les personnes qui entrent les données (y compris les techniciens en informatique). Ces personnes font généralement partie du personnel de l'organisation qui a officiellement la responsabilité de recueillir ces données. Leur expérience passée de la collecte de données dans le cadre d'enquêtes sur les ménages est indispensable. Ces personnes sont les mieux à même de savoir les types de questions auxquels peuvent répondre les familles et les types auxquels elles ne peuvent répondre. Ce groupe devrait comprendre quelqu'un qui ait l'expérience de la phase d'entrée des données du processus de collecte de données. De simples suggestions formulées par quelqu'un possédant une telle expérience peuvent accroître sensiblement la précision des données recueillies et réduire le temps nécessaire pour préparer les données pour l'analyse.

2. Élaboration de la première version du questionnaire

42. La première version de presque tous les questionnaires d'enquête sur les ménages est le fruit d'une série de réunions des membres de l'équipe chargée de l'enquête. Comme pour tous les premiers jets, ce produit ne peut manquer de contenir de nombreuses erreurs. L'approche par modules préconisée dans ce chapitre implique que ce premier jet se compose d'un ensemble de modules différents. Lorsque l'on réunit ces modules dans la première version du questionnaire, il y a plusieurs choses à vérifier.

43. Premièrement, l'équipe chargée de l'enquête doit s'assurer que les modules permettent collectivement de recueillir toutes les informations désirées. Il peut arriver que l'on suppose qu'une question clé pour un module figure dans un autre module,

alors qu'il n'en est rien. Il faut une réunion de tous les participants travaillant sur tous les modules pour s'assurer que certaines questions importantes n'ont pas été omises. De même, il faut éviter les redondances. Lorsque tous les modules sont réunis, il peut arriver que certaines questions aient été posées dans deux différents modules. Ces redondances sont à éliminer pour ne pas faire perdre leur temps aux déclarants et aux enquêteurs. Le seul cas où une même question posée à deux endroits ne doit pas être éliminée à aucun des deux est lorsqu'elle donne confirmation d'une information très importante, par exemple si elle confirme qu'une personne fait vraiment partie de la famille. L'âge des membres de la famille peut être vérifié par des questions sur leur âge actuel et sur leur date de naissance, et le fait qu'une personne fait vraiment partie de la famille peut se vérifier en lui demandant si il ou elle a vécu en d'autres endroits au cours des 12 derniers mois et, si tel est le cas, combien de mois (après lui avoir déjà demandé depuis combien de mois il/elle vit dans la famille interrogée).

44. Deuxièmement, il faut vérifier la longueur totale du questionnaire. Dans n'importe quel pays, il y a une limite au temps que les déclarants sont disposés à consacrer à répondre aux questions d'une enquête sur les ménages. En même temps, les concepteurs ont tendance à vouloir poser un grand nombre de questions, ce qui rend le produit fini beaucoup plus long que prévu à l'origine. L'essai sur le terrain (examiné ci-après) peut aider à répondre à la question de savoir combien de temps il faut pour interroger un ménage type (et combien de temps les déclarants sont disposés à consacrer à leur entrevue), mais les enquêteurs et superviseurs expérimentés peuvent donner aux membres de l'équipe une réponse approximative en examinant le questionnaire. L'élimination des questions visant à recueillir des informations moins prioritaires est un élément douloureux mais nécessaire de l'élaboration de la première version de tout questionnaire d'enquête sur les ménages.

45. Enfin, il faut vérifier la première version du questionnaire pour s'assurer de l'uniformité des périodes de réminiscence. Par exemple, l'un des buts de l'enquête peut être de recueillir des informations sur toutes les sources de revenu du ménage au cours du mois écoulé ou de l'année écoulée⁸. Il faut vérifier le questionnaire afin de s'assurer que toutes les sections où sont recueillies des données sur les revenus portent sur la même période. La principale exception à cette règle tient aux cas occasionnels où, comme on l'a expliqué précédemment, il faut laisser aux déclarants assez de latitude dans le choix de la période de réminiscence qu'il leur est le plus facile d'utiliser.

⁸ Certaines enquêtes font référence à des périodes passées, par exemple en posant des questions sur la situation 5 ou 10 ans auparavant. Il faut aussi vérifier sur tout le questionnaire l'homogénéité de ces points de référence, qui comprennent parfois une date précise, un mois ou une année.

3. Essais sur le terrain et établissement de la version finale du questionnaire

46. Aucun questionnaire d'enquête sur les ménages, si court ou si simple soit-il, ne peut être arrêté dans sa version finale sans avoir été d'abord testé sur un petit nombre de ménages, afin d'identifier ses éventuels défauts de conception. Dans presque tous les cas, un nouveau questionnaire comprend de nombreuses erreurs et insuffisances qui n'apparaissent que lorsqu'il est essayé sur certains ménages types de la population à étudier. On trouvera ci-après quelques règles d'ordre général; pour un examen plus détaillé, voir Grosh et Glewwe (2000) et Converse et Presser (1986).

47. L'expérimentation sur le terrain du projet de questionnaire se déroule en deux stades. Le premier, souvent appelé essai préliminaire, consiste à tester certaines sections (modules) du questionnaire sur un petit nombre de ménages (par exemple, 10-15), pour se faire une idée approximative de sa qualité. Cet essai peut être répété plusieurs fois, en commençant au début du processus de conception du questionnaire. Le second stade est un essai complet du projet de questionnaire sur le terrain. On l'appelle souvent essai pilote. C'est une opération de plus grande envergure, qui porte

sur 100 à 200 ménages, qui doivent être choisis non pas dans une zone limitée mais dans plusieurs zones représentatives de la population à étudier. Pour les enquêtes à effectuer à la fois dans les zones urbaines et rurales, l'essai pilote doit se dérouler dans les unes comme dans les autres. Il doit s'effectuer dans différentes parties du pays ou de la région où la version finale du questionnaire sera utilisée. Enfin, le choix des ménages doit être tel que tous les modules sont testés sur au moins 50 ménages – et de préférence sur plus. Cela veut dire, par exemple, que si le questionnaire comprend un module de collecte de données sur les petites entreprises familiales, au moins 50 des ménages interrogés pour l'essai pilote doivent avoir une telle entreprise.

48. La plupart des essais pilotes nécessitent une période d'une à deux semaines pour interroger 100 à 200 ménages. Tous les membres de l'équipe de l'enquête doivent participer à l'essai pilote et assister à autant d'entrevues que possibles. En fait, les essais pilotes offrent un excellent moyen d'acquérir de l'expérience à quiconque a peu d'expérience de la conception de questionnaires d'enquête sur les ménages. L'une des informations importantes que fournit l'essai sur le terrain est une estimation du temps nécessaire pour remplir le questionnaire⁹. Cependant, il faut aussi se rendre compte que la durée obtenue tend à surestimer (parfois dans un rapport du simple au double) le temps nécessaire pour interroger un ménage dans le cadre de l'enquête proprement dite, d'abord parce que les enquêteurs qui effectuent l'essai pilote ont peu d'expérience du questionnaire, et ensuite parce qu'ils sont ralentis par les défauts du projet de questionnaire qui seront corrigés dans la version finale de celui-ci.

49. L'autre question clé est que dans les pays où l'on parle plusieurs langues, le questionnaire doit être traduit dans toutes les langues importantes et que l'essai pilote doit être effectué dans toutes ces langues. Cela est très important. En particulier, il convient d'éviter autant que possible que les enquêteurs aient à traduire pendant l'entrevue les questions de la langue dans laquelle elles sont rédigées dans celle que parle le déclarant. Des études ont montré (par exemple, Scott et *al.*, 1988) qu'une telle traduction sur place, comparée à une traduction préalable du questionnaire dans la langue du déclarant augmente de deux à quatre fois le risque d'erreurs. Pour vérifier l'exactitude d'une traduction, une personne ou un groupe de personnes autres que les auteurs de la traduction devraient « retraduire » le questionnaire dans sa langue originale. Cette nouvelle traduction serait alors comparée avec l'original, afin de déterminer si elle en reflète fidèlement la teneur, et toute différence signifierait que quelque chose aurait été « perdu dans la traduction ». Pour plus d'informations sur la question de la traduction du questionnaire, voir Harkness, Van de Vijver et Mohler (2003).

50. La dernière considération importante concernant l'essai pilote est qu'il doit servir à tester non seulement le projet de questionnaire mais l'ensemble du plan de travail de terrain, y compris les méthodes de supervision, l'entrée des données et les documents, tels que les manuels à l'usage des enquêteurs (tout cela est examiné de façon plus approfondie au chapitre IV). Ce n'est qu'en testant l'ensemble du processus que l'équipe peut s'assurer que l'enquête est prête à être exécutée. Il peut être utile d'effectuer une « brève analyse » des données recueillies lors de l'essai pilote afin de dépister les éventuels problèmes qui auraient pu échapper à l'attention.

51. Immédiatement après l'essai pilote, l'équipe de l'enquête devrait tenir plusieurs journées de réunions afin de discuter de ses résultats et de modifier le questionnaire à la lumière des leçons tirées. La brève analyse des données de l'essai pilote mentionnée au paragraphe précédent, qui se présente habituellement sous la forme de simples tableaux, devrait être préparée pour ces réunions. Dans certains cas, il peut se poser tant de problèmes qu'un second essai pilote, peut-être de moins grande envergure que le premier, doit être prévu pour vérifier si d'importantes modifications du

⁹ Lors de la conduite des essais préliminaires et des essais pilotes, il conviendrait de prévoir un espace sur le projet de questionnaire pour inscrire l'heure à laquelle on a commencé et l'heure à laquelle on a terminé de remplir chaque module du questionnaire ; ces heures doivent être enregistrées pour chaque ménage interrogé, afin d'indiquer le temps qu'il faut pour remplir chaque module.

questionnaire donneront de bons résultats sur le terrain. Tous les membres de l'équipe doivent être présents à ces réunions, auxquelles devraient également participer la plupart sinon la totalité des personnes ayant effectué les entrevues pendant l'essai pilote.

52. Un volume considérable de travaux de recherche ont été effectués ces dernières années sur la conception du questionnaire, et de nouvelles méthodes précieuses ont été élaborées pour la conception de questionnaires efficaces. Si ces méthodes ne sont pas encore amplement utilisées dans les pays en développement et les pays en transition, leur utilisation a des chances d'augmenter très sensiblement à l'avenir. Nous n'avons pas la place pour décrire ici ces méthodes, mais nous encourageons le lecteur à consulter les écrits qui leur ont été consacrés. Ces méthodes comprennent des groupes de réflexion, des entrevues cognitives et la codification des comportements. Esposito et Rothgeb (1997), Biemer et Lyberg (2003) donnent un bon aperçu de ces méthodes. Voir aussi Krueger et Casey (2000) pour les groupes de réflexion, Forsyth et Lessler (1991) pour les entrevues cognitives, et Fowler et Cannell (1996) pour la codification des comportements. On trouvera également au chapitre IX de cette publication des informations sur les groupes de réflexion et la codification des comportements, respectivement aux sections C.2 et C.6.

E. Conclusion

53. Ce chapitre a présenté des recommandations générales pour la conception de questionnaires d'enquêtes sur les ménages pour les pays en développement. Il a porté plus spécialement sur les questionnaires administrés aux ménages. Certaines enquêtes sur les ménages servent également à recueillir des données sur la communauté locale, à l'aide d'un « questionnaire communautaire » distinct. Voir Frankenberg (2000) pour des recommandations détaillées sur la conception de ce type de questionnaires.

54. S'il a traité de nombreux sujets, chaque sujet n'a été traité que brièvement. Quiconque envisage une telle enquête doit consulter d'autres documents afin d'en tirer des conseils beaucoup plus détaillés. À cet égard, les références indiquées à la fin de ce chapitre offrent un bon point de départ.

RÉFÉRENCES

- Biemer, Paul P., et Lars E. Lyberg (2003). *Introduction to Survey Quality*. New York, Wiley.
- Casley, Dennis, et Denis Lury (1987). *Data Collection in Developing Countries*. Oxford, Royaume-Uni, Clarendon Press.
- Converse, Jean M., et Stanley Presser (1986). *Survey Questions: Handcrafting the Standardized Questionnaire*. Beverly Hills, Californie, Sage Publications.
- Deaton, Angus, et Margaret Grosh (2000). Consumption. Dans *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 Years of the Living Standards Measurement Study*, Margaret Grosh et Paul Glewwe, eds. New York, Oxford University Press (pour la Banque mondiale).
- Esposito, James L., et Jennifer M. Rothgeb (1997). Evaluating survey data: making the transition from pretesting to quality assessment. Dans *Survey Measurement and Process Quality*, Lars E. Lyberg *et al.*, eds. New York, Wiley.
- Forsyth, Barbara H., et Judith T. Lessler (1991). Cognitive laboratory methods: a taxonomy. Dans *Measurement Errors in Surveys*, Paul P. Biemer *et al.*, eds. New York, Wiley.

- Fowler, F.J., et C.F. Cannell (1996). Using behavior coding to identify cognitive problems with survey questions. dans *Methodology for Determining Cognitive and Communicative Processes in Survey Research*. San Francisco, Californie, Jossey-Bass.
- Frankenberg, Elizabeth (2000). Community and price data. Dans *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 Years of the Living Standards Measurement Study*, Margaret Grosh et Paul Glewwe, eds. New York, Oxford University Press (pour la Banque mondiale).
- Grosh, Margaret, et Paul Glewwe, eds. (2000). *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 Years of the Living Standards Measurement Study*. New York, Oxford University Press (pour la Banque mondiale).
- Harkness, Janet A., Fons J. R. Van de Vijver et Peter Mohler (2003). *Cross-Cultural Survey Methods*. New York, Wiley
- Husmanns, R., F. Merhan et V. Verma (1990). *Surveys of Economically Active Population, Employment, Unemployment, and Underemployment*. An ILO Manual on Concepts and Methods. Genève, Organisation internationale du Travail.
- Krueger, Richard A., et Mary Anne Casey (2000). *Focus Groups: A Practical Guide for Applied Research*. Thousand Oaks, Californie, Sage Publications.
- Nations Unies (1985). *United Nations National Household Survey Capability Programme: Development and Design of Survey Questionnaires* (INT-84-014). New York.
- Nations Unies (1993). *National Household Survey Capability Programme: Sampling Rare and Elusive Populations* (INT-92-P80-16E). New York.
- Scott, Christopher, *et al.* (1988). Verbatim questionnaires versus field translations or schedules: an experimental study. *International Statistical Review*, vol. 56, n° 3, pp. 259-78.
- Sudman, Seymour, et Norman M. Bradburn (1982). *Asking Questions. A Practical Guide to Questionnaire Design*. San Francisco, Californie, Jossey-Bass.

Chapitre IV

Vue d'ensemble de la mise en œuvre des enquêtes sur les ménages dans les pays en développement

PAUL GLEWWE

Département d'économie appliquée
Université du Minnesota
Saint Paul, Minnesota, États-Unis d'Amérique

RÉSUMÉ

Le présent chapitre contient un examen des questions fondamentales concernant la mise en œuvre des enquêtes sur les ménages dans les pays en développement, en commençant par les activités qui doivent être menées avant le lancement de l'enquête sur le terrain : l'établissement d'un budget et d'un plan de travail, la sélection d'un échantillon, la formation du personnel chargé de l'enquête et la conception de manuels didactiques, et l'élaboration du plan pour le travail de terrain. Il traite également des activités qui se déroulent tandis que l'enquête se poursuit : mise en place et maintien de moyens de communication et de transport adéquats, établissement de protocoles de supervision et autres activités destinées à rehausser la qualité des données, et établissement d'un système de gestion des données. Il se termine par une brève section sur les activités menées après achèvement du travail de terrain, suivie d'une brève conclusion.

Termes clés : mise en œuvre de l'enquête, budget, plan de travail, échantillon, formation, plan pour le travail de terrain, communications, transports, supervision, gestion des données.

A. Introduction

1. La valeur des informations que fournissent les enquêtes sur les ménages est étroitement liée à l'utilité et à l'exactitude des données recueillies, qui dépendent elles-mêmes de la façon dont l'enquête est menée sur le terrain. Le présent chapitre contient des recommandations générales sur l'exécution des enquêtes, qui portent sur presque tous les aspects de leur processus d'exécution, hormis la conception du questionnaire.

2. On peut considérer qu'un questionnaire (y compris le plan d'analyse des données correspondant) bien conçu représente la moitié du parcours nécessaire au succès d'une enquête. La destination finale est atteinte par un déroulement efficace de l'enquête. Celui-ci est amorcé non pas lorsque les enquêteurs commencent à interroger les ménages qui leur sont affectés mais des mois, et souvent un ou deux ans, plus tôt. La section B de ce chapitre présente un examen des tâches à accomplir avant

que les ménages puissent être interrogés; la section C décrit les activités qui sont menées tandis que l'enquête s'effectue sur le terrain; la section D offre une brève analyse des tâches à accomplir une fois le travail de terrain terminé; et la section finale offre quelques brèves remarques à titre de conclusion. Si ce chapitre offre une présentation utile de ce sujet, il est beaucoup trop bref pour contenir un exposé détaillé de tous les conseils nécessaires. Pour s'assurer que l'enquête atteindra ses objectifs, ses auteurs feraient bien de consulter des ouvrages beaucoup plus approfondis. L'étude de Grosh et Muñoz (1996) est un bon point de départ : bien qu'elle se concentre sur les études de la Banque mondiale sur la mesure du niveau de vie, une large part de ses conseils sont applicables à presque tous les types d'enquêtes sur les ménages. Les deux autres références utiles sont Casley et Lury (1987) et les Nations Unies (1984).

3. Tout au long de ce chapitre, on suppose que l'enquête est préparée et exécutée par une équipe principale bien organisée désignée à cet effet. On suppose également que le questionnaire est utilisé par des enquêteurs qui rendront visite aux déclarants dans leurs foyers et que l'unité d'échantillonnage est le ménage¹. Enfin, le lecteur ne doit pas perdre de vue que ce chapitre se concentre sur les pays en développement, y compris les économies à faible revenu en transition comme la Chine et le Viet Nam. Néanmoins, la plupart des recommandations s'appliquent aussi aux économies en transition plus développées de l'Europe de l'Est et de l'ex-Union soviétique.

¹ Pour certaines enquêtes, l'unité d'échantillonnage est le logement et non pas le ménage, mais en pareil cas certains ou tous les ménages des logements de l'échantillon deviennent les « unités déclarantes » de l'enquête.

B. Tâches à accomplir avant que l'équipe de l'enquête se rende sur le terrain

4. Pour n'importe quelle enquête sur les ménages, la première tâche est de former une équipe centrale qui sera chargée de gérer tous les aspects de l'enquête. La chapitre III a indiqué de façon détaillée qui devrait faire partie de cette équipe. Une fois que l'équipe centrale est en place, elle doit s'acquitter des huit tâches suivantes avant qu'aucun ménage puisse être interrogé :

- a) Établir un budget provisoire et mobiliser les fonds correspondants;
- b) Élaborer un plan de travail pour toutes les autres tâches;
- c) Choisir un échantillon de ménages à interroger;
- d) Concevoir des manuels de formation;
- e) Former le personnel de terrain et les agents chargés d'entrer les données;
- f) Préparer un plan pour le travail de terrain et l'entrée des données;
- g) Procéder à un essai pilote;
- h) Lancer une campagne de publicité.

Cette liste de tâches est présentée sensiblement en ordre chronologique. Chaque tâche est décrite ci-dessous.

1. Financement du budget

5. Les ressources financières limitent sérieusement ce qui peut être accompli par toute enquête sur les ménages. Dans presque tous les cas, la première tâche est d'établir un projet de budget fondé sur diverses hypothèses quant au nombre de ménages à choisir pour l'échantillon et au temps nécessaire pour interroger un ménage type. Ce budget ne peut être qu'approximatif car certains détails de coût ne peuvent être connus que lorsque l'on a pris connaissance du questionnaire, mais le plus souvent, le budget final est assez proche du projet de budget (à moins que les objectifs de l'enquête n'aient été sensiblement modifiés).

6. Une fois le projet de budget préparé, il faut réunir les fonds nécessaires. Si le financement est incertain, il est préférable d'attendre qu'il ait été assuré avant d'entreprendre la planification détaillée de l'enquête. Cela éviterait un gaspillage de temps et de personnel, au cas où les crédits ne pourraient être obtenus.

7. Bien qu'il soit difficile d'en dire beaucoup plus sur l'établissement du budget sans en savoir davantage sur la nature et le type de l'enquête, on peut faire quelques recommandations générales. Premièrement, il conviendrait d'évaluer la capacité de l'organisation chargée d'effectuer l'enquête. Si cette organisation ne dispose pas de certaines compétences techniques — par exemple, si elle a peu de personnel qualifié en matière d'établissement d'échantillons ou d'utilisation des nouvelles technologies de l'information —, il peut être nécessaire qu'elle fasse appel à des consultants extérieurs. Cela pourrait alourdir considérablement le coût de l'enquête, quoique le plus souvent, cet excédent de coût soit clairement justifié. Deuxièmement, il serait bon de commencer par examiner les budgets d'enquêtes similaires effectuées précédemment dans le pays ou dans des pays comparables. Troisièmement, pour éviter la gêne imposée par les coûts imprévus, un « coussin » d'environ 10 % du budget total devrait être explicitement ajouté comme poste budgétaire supplémentaire. Ce poste est souvent désigné sous le nom de *provision pour imprévus*. Lorsque l'incertitude concernant les coûts est grande, cette provision peut être de 15 ou même de 20 %.

8. Pour donner une vision plus concrète de ce qui précède, le tableau IV.1 [version modifiée du tableau 8.2 de l'étude de Grosh et Muñoz (1996)] présente un projet de budget correspondant à un cas hypothétique. Dans cet exemple, on suppose que 3 000 ménages seront interrogés dans le cadre de l'enquête, et que la collecte de données s'étalera sur un an. Outre l'équipe centrale (voir chapitre III), il y a quatre équipes de terrain, composées chacune de trois enquêteurs, d'un superviseur et d'un agent chargé d'entrer les données. Deux conducteurs transporteront les équipes à leur lieu de travail à bord de véhicules affectés au projet. On suppose que chaque enquêteur travaillera 250 jours dans l'année, interrogeant (en moyenne) un ménage par jour. Le tableau IV.1 présente les salaires hypothétiques de tout le personnel, ainsi que d'hypothétiques « indemnités pour frais de déplacement » versées aux membres des équipes pour chaque journée de travail sur le terrain. Chaque équipe de terrain sera munie d'un ordinateur pour l'entrée des données, et l'équipe centrale disposera de trois ordinateurs pour leur analyse. Des coûts hypothétiques sont également indiqués pour les consultants, tant internationaux que locaux. Bien entendu, ce tableau n'est donné qu'à titre indicatif : le coût de toute enquête donnée dépendra de la taille de l'échantillon, du nombre de personnes engagées, de leurs salaires et autres formes de rémunération, du ratio superviseur/enquêteurs, du nombre de ménages qu'un enquêteur peut couvrir en une journée, de la question de savoir si l'entrée des données se fait sur le terrain ou en un lieu central, et de nombreux autres facteurs. Il n'est présenté ici que pour servir de « liste récapitulative », afin de faire en sorte que tous les coûts essentiels soient inclus dans le projet de budget de l'enquête.

2. Plan de travail

9. Une fois que le financement a été obtenu, la tâche suivante consiste à établir un plan de travail réaliste, qui est essentiellement un calendrier des activités depuis les premiers stades de la planification de l'enquête jusqu'à l'achèvement du travail de terrain². Le plan de travail comprend chacune des activités suivantes : gestion d'ensemble (y compris les achats d'équipement); conception du questionnaire; choix de l'échantillon; affectation, recrutement et formation de personnel; entrée et gestion des données; activités de terrain; et analyse des données, traitement, documentation et ré-

² Il s'agit d'un plan de travail général comprenant beaucoup de tâches à accomplir avant que démarre le travail de terrain (c'est-à-dire avant que des ménages soient interrogés). Il faut aussi prévoir un « plan de travail de terrain et un plan d'entrée des données », comme on le verra ci-après.

daction d'un rapport. Pour chacune de ces activités, il faut établir une liste des tâches à accomplir et fixer les dates de leur exécution (autrement dit, des délais d'exécution). Les principales étapes, telles que l'essai pilote et le premier jour de travail de terrain, doivent être indiquées plus particulièrement. Cette liste, qui peut souvent être présentée sous forme de tableau, est le plan de travail de l'enquête.

Tableau IV.1
Projet de budget d'une enquête hypothétique sur 3 000 ménages
(en dollars des États-Unis)

Poste	Nombre	Durée	Coût unitaire	Coût total
Salaires de base				
Chef du projet	1	30 mois	800/mois	24 000
Responsable des données	1	30 mois	600/mois	18 000
Chef des équipes de terrain	1	30 mois	600/mois	18 000
Assistants/comptable	3	24 mois	450/mois	32 400
Superviseurs	4	14 mois	400/mois	22 400
Enquêteurs	12	13 mois	350/mois	54 600
Agents chargés de l'entrée des données	4	13 mois	300/mois	15 600
Chauffeurs	2	13 mois	300/mois	7 800
				Total partiel 192 800
Indemnités de déplacement				
Chef du projet	1	90 jours	30/jour	2 700
Responsable des données	1	60 jours	30/jour	1 800
Chef des équipes de terrain	1	90 jours	30/jour	2 700
Assistants	2	60 jours	30/jour	3 600
Personnel chargé des listes	10	60 jours	15/jour	9 000
Superviseurs	4	290 jours	15/jour	17 400
Enquêteurs	12	270 jours	15/jour	48 600
Chauffeurs	2	270 jours	15/jour	8 100
				Total partiel 93 900
Matériels				
Achat de véhicules	2	—	20 000	40 000
Carburant et entretien	2	13 mois	300/mois	7 800
Ordinateurs (entrée des données)	4	—	1 000	4 000
Imprimantes, stabilisateurs, etc.	5	—	1 000	5 000
Ordinateurs (analyse des données)	3	—	1 500	4 500
Informatique/fournitures de bureau	—	30 mois	350/mois	10 500
Photocopieur/télécopieur	1	—	2 500	2 500
				Total partiel 74 300
Coûts d'impression				
Questionnaires	3 500	—	2	7 000
Manuels de formation	40	—	5	200
Rapports	500	—	5	2 500
				Total partiel 9 700

Poste	Nombre	Durée	Coût unitaire	Coût total
Coûts des consultants				
Consultants étrangers	5	Personnes/mois	10 000/mois	50 000
Indemnités journalières (personnel international)	150	Jours	150/jour	22 500
Voyages internationaux	8	Voyages	2 000/voyage	16 000
Consultants locaux	5	Personnes—mois	3 000/mois	15 000
Total partiel				103 500
Imprévu (10 %)				47 400
Total				521 600

Note : Un tiret (—) signifie « sans objet ».

10. Bien entendu, un grand nombre de ces activités sont étroitement liées et doivent donc être coordonnées. Par exemple, de nombreuses activités de gestion des données et d'analyse des données ne peuvent commencer avant que l'équipement nécessaire ait été acheté et que le personnel auquel elles seront confiées ait été affecté (ou recruté) et formé. Il faut également tenir compte du fait que même les meilleurs plans ont besoin d'être modifiés lorsque se produisent des imprévus. À posteriori, la plupart des plans sont trop optimistes, et les retards sont fréquents. Dans la mesure du possible, le calendrier des diverses activités doit être réaliste et doit comprendre des « temps de pause » destinés à permettre aux participants de rattraper les retards qui ne peuvent manquer de se produire.

11. La figure IV.1 [adaptée de la figure 8.2 de l'étude de Grosh et Muñoz (1996)] offre un exemple de plan de travail. Ce plan porte sur une période de 30 mois. Les astérisques (*) indiquent les moments où diverses activités se déroulent simultanément. Cette figure montre que les préparatifs doivent commencer environ un an avant que l'enquête démarre sur le terrain. Le fait que l'essai pilote se produit dans le huitième mois implique qu'un projet de questionnaire, un personnel formé et un projet de programme d'entrée des données soient prêts à ce stade. Le travail de terrain doit débiter le douzième mois et se poursuivre pendant un an. Le plan de travail suppose également qu'un projet de rapport soit rédigé lorsque la moitié des données a été recueillie. Bien entendu, les plans de travail d'une enquête donnée ne peuvent manquer de différer de celui-ci. Ce modèle doit servir de liste récapitulative et indiquer l'échelonnement dans le temps des différentes tâches à coordonner.

3. Sélection d'un échantillon de ménages

12. Dans la plupart des enquêtes sur les ménages, il y a une population à étudier, par exemple la population de tout le pays représentée par les ménages de l'enquête. Le processus de sélection d'un ensemble de ménages représentatifs de la population dans son ensemble est appelé échantillonnage, et la procédure d'établissement de l'échantillon est appelée conception de l'échantillon. Lors de l'établissement de l'échantillon, il faut tenir compte d'un grand nombre de questions, si grand qu'il n'est même pas possible de les énumérer toutes dans un aperçu aussi bref que celui-ci. Voir chapitres II, V et VI de ce volume pour des recommandations détaillées sur l'échantillonnage. Une initiation à l'échantillonnage est offerte par Kalton (1983); des analyses beaucoup plus détaillées en sont données par Kish (1965), Cochran (1977) et Lohr (1999).

13. Les commentaires sur l'échantillonnage présentés dans ce chapitre se limiteront à deux remarques à prendre en compte par l'équipe de l'enquête. Premièrement, il est parfois utile de concevoir l'échantillon de manière que les ménages soient interrogés sur une période de 12 mois. Cela permet de compenser les variations saisonnières des phénomènes étudiés, et aussi d'utiliser les données pour étudier les comportements saisonniers. Le deuxième point, plus important encore, c'est que les responsables de la planification de l'enquête doivent résister à la tentation d'inclure un grand nombre de ménages dans l'échantillon. Il est naturel qu'ils cherchent à établir un échantillon aussi grand que possible, surtout pour les groupes auxquels ils s'intéressent plus particulièrement, car cela réduit le risque d'erreur d'échantillonnage. Toutefois, dans bien des cas, l'accroissement de la taille de l'échantillon s'accompagne d'un accroissement du nombre d'erreurs non liées à l'échantillonnage, telles que les erreurs dues à l'utilisation d'un personnel moins qualifié et à des ratios superviseur/enquêteurs plus faibles. Il est très possible, et peut-être même probable, que l'accroissement du nombre d'erreurs non liées à l'échantillonnage pèse plus lourd que les réductions des erreurs d'échantillonnage dues à l'utilisation d'un plus gros échantillon.

Figure IV.1
Plan de travail pour l'élaboration et l'exécution
d'une enquête sur les ménages

Tâche	Mois de l'enquête																														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	
Gestion et logistique																															
Constitution de l'équipe centrale	*																														
Achat d'ordinateurs		*				*	*	*	*																						
Achat de matériels pour l'enquête						*	*	*																							
Publicité							*	*	*	*	*	*																			
Achat/location de véhicules					*	*	*	*	*																						
Élaboration du questionnaire																															
Définition des objectifs de l'enquête		*	*																												
Préparation du projet de questionnaire			*	*	*																										
Réunions sur le projet de questionnaire						*	*																								
Version finale du projet de questionnaire pour l'essai pilote							*																								
Essai pilote								*																							
Réunions après l'essai pilote									*																						
Impression de la version finale du questionnaire										*																					
Échantillonnage																															
Conception de l'échantillon et cadre d'échantillonnage		*	*																												
Sélection de l'échantillon (UPE)				*																											
Établissement du plan de travail de terrain				*																											
Listes/mappage des UPE					*	*	*	*	*																						
Recrutement et formation																															

Tâche	Mois de l'enquête																													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Sélection et formation du personnel chargé de l'essai pilote						*	*																							
Préparation des manuels de formation									*	*																				
Formation des enquêteurs										*																				
Gestion des données																														
Conception du premier programme d'entrée des données						*	*	*																						
Version finale du programme d'entrée des données									*	*																				
Rédaction du manuel d'entrée des données									*	*																				
Formation du personnel chargé d'entrer les données										*																				
Travail de terrain																														
										*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Analyse et documentation																														
Rédaction du plan d'analyse																		*	*											
Analyse de la première moitié des données																			*	*										
Rédaction d'un rapport préliminaire																					*									
Création d'un premier ensemble complet de données																								*	*					
Première analyse des données																										*	*			
Rapport final et documentation																												*	*	*

4. Rédaction des manuels de formation

14. L'élément le plus important de la formation est peut-être la préparation de manuels pour toutes les personnes qui recevront une formation d'enquêteur, de superviseur et d'agent chargé de l'entrée des données. Il faut des manuels distincts pour chaque catégorie, soit un manuel de l'enquêteur, un manuel du superviseur et un manuel d'entrée des données. Ces manuels sont d'une importance critique pour la formation et doivent être prêts avant que celle-ci ne commence. Mais surtout, ils doivent servir de documents de référence lorsque l'enquête elle-même est en cours et fournir toutes les informations nécessaires aux différentes catégories de personnel de terrain en même temps qu'aux agents chargés de l'entrée des données³. En effet, les analystes des données les utilisent souvent pour mieux comprendre les données qu'ils analysent, ce qui implique que des exemplaires supplémentaires de tous les manuels soient produits à leur intention. En règle générale, chaque fois qu'il peut y avoir un doute sur l'utilité d'une information, il vaut mieux l'inclure dans le manuel que de ne pas l'y mentionner.

15. Les manuels de formation doivent expliquer le but de l'enquête et les tâches élémentaires à effectuer par le personnel auxquels ils sont destinés. Les procédures à utiliser pour les cas inhabituels doivent également y figurer, y compris les principes généraux à appliquer face aux problèmes imprévus. Les manuels doivent également expliquer comment remplir les formulaires prévus (cela est particulièrement important pour le manuel du superviseur). Dans la mesure où même les manuels les mieux

³ L'expression « agent de terrain » s'applique aux enquêteurs, superviseurs et autres agents qui, pour effectuer leur travail, doivent se rendre dans les communautés où des ménages sont interrogés. Comme il est indiqué ci-dessous, il est très utile de rapprocher le plus possible de ces communautés les agents chargés de l'entrée des données. Dans le cas des enquêtes pour lesquelles ces agents se déplacent avec le personnel de terrain, ils sont aussi désignés comme agents de terrain, mais pas dans les autres cas. Dans ce chapitre, l'expression « agents de terrain et agents d'entrée des données » s'applique dans ces deux cas.

conçus peuvent présenter des erreurs ou omissions, il conviendrait de préparer un ou plusieurs jeux « d'instructions supplémentaires » comme suppléments à ces manuels après qu'ils ont été distribués au personnel de terrain et aux agents chargés de l'entrée des données.

5. Formation du personnel de terrain et des agents chargés de l'entrée des données

16. Parfois, l'organisation qui effectue l'enquête dispose d'un grand nombre d'enquêteurs, de superviseurs et d'agents d'entrée des données expérimentés. Lorsque l'enquête ressemble beaucoup à d'autres effectuées précédemment par cette organisation, le temps de formation peut être court, d'une semaine ou deux, juste assez pour expliquer les détails du nouveau questionnaire et certains changements de méthode introduits pour la nouvelle enquête. Dans d'autres cas, en revanche, cette nouvelle enquête peut être très différente de toutes celles que l'organisation a pu effectuer récemment, et celle-ci doit alors engager au moins quelques nouveaux agents de terrain et d'entrée des données. Dans ces cas, il faut dispenser à ceux-ci une formation approfondie pour assurer la qualité de l'enquête. Par exemple, les enquêteurs et superviseurs nouvellement engagés doivent recevoir une formation générale avant d'être initiés aux particularités de la nouvelle enquête. En général, il faut alors plus de deux semaines de formation, plutôt trois ou quatre, pour que les enquêteurs et les superviseurs soient prêts à faire un travail efficace.

17. Bien que la nature de la formation soit fonction de celle de l'enquête, on peut quand même faire quelques observations générales. Premièrement, la formation doit comprendre une grande quantité de travaux pratiques, comportant l'utilisation du questionnaire pour interroger véritablement des ménages. Deuxièmement, elle doit mettre l'accent sur la compréhension des objectifs de l'enquête et de la façon dont les données recueillies contribueront à leur réalisation. En insistant sur cette compréhension, plutôt que d'inciter les agents de terrain et d'entrée des données à suivre strictement des règles prescrites, on permet aux enquêteurs et à leurs superviseurs de faire face aux problèmes imprévus qui peuvent surgir. Troisièmement, il vaut mieux former plus de personnes qu'il n'est nécessaire et les soumettre à une certaine forme d'examen (avec « écrit » et « oral pratique »). Les résultats de cet examen peuvent alors être utilisés pour sélectionner comme enquêteurs et superviseurs ceux qui ont obtenu les meilleurs résultats. Quatrièmement, la formation devrait être centralisée, afin que tous les agents de terrain reçoivent la même, et elle devrait être de grande qualité. Enfin, il faut bien comprendre que la qualité de la formation peut être déterminante pour la qualité de l'enquête et, en fin de compte, la qualité des données recueillies. L'équipe de l'enquête tout entière doit accorder une grande attention à la formation et ne pas se contenter de déléguer celle-ci à un ou deux de ses membres.

6. Travail de terrain et plan d'entrée des données

18. Le travail effectif qui consiste à se rendre dans les zones étudiées et à interroger les ménages de l'échantillon est généralement appelé le travail de terrain. Comme ce travail doit être étroitement coordonné avec l'entrée des données, ces deux activités sont examinées ensemble dans ce chapitre. Le travail de terrain doit débiter le plus tôt possible (parfois moins d'une semaine) après la formation, afin de réduire les risques d'oubli de ce qui a été appris. Avant que ce travail ne puisse commencer, il convient d'établir un plan très détaillé affectant chaque ménage sélectionné (à partir du plan d'échantillonnage) à tel enquêteur, superviseur et agent d'entrée des données.

Chaque équipe est chargée d'une fraction de l'échantillon total et doit s'assurer que les ménages qui lui ont été attribués sont bien interrogés.

19. Lors de l'élaboration du plan de travail de terrain, il faut garder plusieurs principes présents à l'esprit. Premièrement, il faut prévoir des moyens de transport adéquats, non seulement pour le personnel mais aussi pour les fournitures. L'expérience des enquêtes sur les ménages effectuées dans de nombreux pays montre que les difficultés logistiques les plus courantes sont liées à l'approvisionnement en carburant et en lubrifiant et à l'entretien des véhicules utilisés par le personnel de terrain. Deuxièmement, le plan de travail de terrain doit être réaliste et s'inspirer de l'expérience des enquêtes précédentes sur les ménages effectuées dans le pays (voir les commentaires sur l'essai pilote au chapitre III). Troisièmement, le plan de travail de terrain devrait s'accompagner d'un plan d'entrée des données expliquant le processus par lequel les informations contenues dans les questionnaires remplis sont entrées sur ordinateur et classées dans des fichiers principaux au bureau central. Quatrièmement, pour les enquêtes qui nécessitent un travail de terrain de plusieurs mois, il faudrait prévoir une pause après les premières semaines afin d'évaluer comment se déroulent le travail de terrain et l'entrée des données⁴. Il est très probable que l'expérience de ces premières semaines débouche sur des suggestions de modifications à apporter au travail de terrain et aux procédures d'entrée des données; ces modifications devraient être enregistrées et communiquées au personnel de terrain sous la forme « d'ajouts » à leurs manuels, comme on l'a expliqué précédemment. Cinquièmement, avant que le plan de travail de terrain ne soit finalisé, il devrait être soumis aux commentaires et suggestions de superviseurs et d'enquêteurs expérimentés. Enfin, il faudrait laisser aux enquêteurs assez de temps auprès de chaque unité primaire d'échantillonnage (UPE) pour qu'ils rendent plusieurs visites aux ménages de l'échantillon afin de s'assurer que les données recueillies émanent des déclarants les mieux informés; l'autre solution, qui consiste à obtenir des « réponses de substitution » d'un autre membre de la famille moins bien informé, risque de réduire l'exactitude des données recueillies.

⁴ Cette pause devrait être observée durant une période normale, afin que la collecte des données ne soit pas interrompue au milieu d'un événement important devant être couvert par l'enquête.

7. Exécution d'un essai pilote

20. Toutes les enquêtes sur les ménages devraient prévoir la conduite d'un « essai » sur le questionnaire, les plans de travail de terrain et d'entrée des données, et sur tous les autres aspects de l'enquête. C'est ce que l'on appelle l'essai pilote. Il comprend des entrevues avec 100 à 200 ménages de toutes les régions du pays qui seront couvertes par l'enquête. Comme l'un des principaux buts de l'essai pilote est d'évaluer la conception du questionnaire, cette question est traitée de façon détaillée au chapitre III. Lorsque l'essai pilote est terminé, une réunion de plusieurs jours est convoquée, au cours de laquelle l'équipe centrale de l'enquête et les participants à l'essai pilote examinent les problèmes que cet essai a permis d'identifier. Les participants à cette réunion doivent alors se mettre d'accord sur la version finale du questionnaire, des plans de travail de terrain et d'entrée des données et sur tous autres aspects de l'enquête.

8. Lancement d'une campagne de publicité

21. Le lancement d'une nouvelle enquête sur les ménages devraient faire l'objet d'une publicité dans les médias, afin d'informer le public et, si possible, d'encourager les ménages choisis pour les entrevues à coopérer. L'autre avantage des campagnes de publicité est qu'elles relèvent le moral du personnel de l'enquête. En général, il n'est pas judicieux de consacrer un gros budget à la publicité auprès du grand public car l'immense majorité des ménages informés ne seront pas interrogés dans le cadre de l'enquête. Toutefois, dans certains cas, on peut obtenir cette publicité presque gratui-

tement en contactant les stations de télévision et de radio, la presse écrite et d'autres organismes des médias. Les articles de journaux sont particulièrement utiles car les enquêteurs et les superviseurs peuvent en conserver des exemplaires à montrer à tout ménage qui aurait des doutes sur ce que les enquêteurs disent au sujet de l'enquête.

22. Une publicité étroitement ciblée est également utile. Il peut s'agir de prospectus affichés dans les communautés choisies comme UPE, ainsi que de lettres aux ménages qui ont été choisis pour être interrogés. Les prospectus affichés devraient être colorés et de nature à attirer le regard, et les lettres comme les prospectus devraient faire ressortir l'utilité des données pour l'amélioration des décisions des pouvoirs publics. Les lettres devraient également souligner que les données recueillies seront traitées comme strictement confidentielles; dans beaucoup de pays, des lois particulières peuvent être invoquées comme garanties de confidentialité. Enfin, il conviendrait de contacter les dirigeants des communautés locales afin de leur expliquer l'importance et les avantages à attendre de l'enquête. Une fois convaincus de ces avantages, ces dirigeants peuvent alors persuader les ménages réticents de participer à l'enquête.

C. Activités de soutien au travail de terrain

23. Une fois que tous les travaux préparatoires sont terminés, les entrevues avec les ménages peuvent commencer. Chaque pays a sa propre manière de mener les enquêtes sur les ménages. Néanmoins, quelques conseils d'ordre général pourraient être utiles à tous (voir immédiatement ci-dessous). On suppose que le travail de terrain est effectué par des équipes mobiles.

1. Communications et transports

24. Chaque équipe a besoin d'un accès à une ligne de communication fiable avec l'administration centrale de l'enquête pour pouvoir lui rendre compte des progrès réalisés et des difficultés rencontrées et pour transmettre le plus vite possible les données recueillies au bureau central. Souvent, les pays en développement ont des moyens de communication insuffisants, surtout dans les régions rurales. Néanmoins, dans la plupart des pays, le service téléphonique s'est amélioré, au point que chaque équipe de terrain peut avoir accès à un téléphone fiable en l'espace de quelques heures ou tout au plus d'un jour ou deux. En effet, les téléphones portables se répandent de plus en plus dans beaucoup de pays en développement, quoique pas toujours dans les zones rurales. La solution simple est de doter ces équipes de téléphones portables fonctionnant dans les régions desservies par cette technologie. Pour les équipes opérant dans des régions reculées, il peut être judicieux d'investir dans des téléphones par satellite.

25. Des transports fiables sont aussi d'une importance capitale pour le travail des équipes de terrain. Le mode de transport employé peut varier d'un pays à l'autre, mais chaque équipe devrait disposer d'un moyen de transport qui lui permette de se rendre sans encombre d'une zone de travail à une autre. L'évacuation d'urgence doit également être prévue pour le cas où un membre d'une équipe de terrain tomberait gravement malade et aurait besoin de soins immédiats. Pour les transports ordinaires et d'urgence, il convient de prévoir un système de secours qui puisse être utilisé en cas de défaillance du système principal. Un moyen de transport fiable peut également servir de mode de communication de rechange si tous les autres moyens échouent.

2. Supervision et assurance de la qualité

26. La qualité du travail accompli par les enquêteurs est d'une importance cruciale pour toute enquête sur les ménages. Assurer la qualité n'est pas chose aisée. Certains enquêteurs peuvent tout simplement se montrer incapables de faire leur travail, et d'autres peuvent ne pas faire tout leur possible s'ils ne sont guère incités à le faire. Le meilleur moyen d'assurer un travail de qualité est de disposer d'un système efficace de supervision du travail de terrain.

27. Les recommandations suivantes devraient aider les superviseurs à se montrer efficaces dans le suivi et le maintien de la qualité du travail des enquêteurs. Premièrement, chaque superviseur ne devrait être responsable que d'un petit nombre d'enquêteurs : pas plus de cinq et, de préférence, deux ou trois. Deuxièmement, chaque superviseur devrait consacrer au moins la moitié de son temps à vérifier la qualité du travail des enquêteurs. Troisièmement, il conviendrait d'établir une liste récapitulative de contrôle relativement brève afin d'aider les superviseurs à vérifier les questionnaires remplis qui leur sont soumis par les enquêteurs; cela leur permettrait de s'assurer que ceux-ci appliquent certaines règles élémentaires lors de leur entrevue avec chaque ménage. *Chaque* questionnaire devrait être vérifié à l'aide de cette liste, et ces contrôles devraient donner lieu à un compte rendu écrit. Quatrièmement, les superviseurs devraient rendre des visites *surprises* aux enquêteurs afin de les observer en action. Cela leur permettrait de s'assurer qu'ils sont là où ils doivent être. De plus, le superviseur devrait observer l'enquêteur tandis qu'il interroge un ménage, afin de vérifier qu'il suit bien toutes les procédures qui lui ont été enseignées lors de la formation. Cinquièmement, les enquêteurs devraient choisir au hasard des ménages auxquels rendre de nouveau visite après leur entrevue avec l'enquêteur. Une autre liste récapitulative plus détaillée devrait être préparée pour la conduite de « mini-entrevues » portant sur certaines questions clés (par exemple, combien de personnes vivent actuellement au sein du ménage), afin de permettre au superviseur de s'assurer que l'enquêteur a bien porté les informations les plus élémentaires sur le questionnaire. Sixièmement, dans le cas d'équipes mobiles, le plan de travail de terrain devrait être conçu de telle manière que le superviseur accompagne les enquêteurs dans leurs déplacements; après tout, le superviseur ne peut superviser comme il faut les enquêteurs s'il est loin d'eux lorsqu'ils effectuent leurs entrevues.

28. On peut faire deux autres recommandations concernant la supervision et le contrôle de la qualité des données. Premièrement, il conviendrait d'envisager sérieusement d'entrer les données sur le terrain à l'aide d'ordinateurs portables et d'utiliser un logiciel qui permette de vérifier la cohérence des données. Toute incongruité constatée peut alors être résolue par une nouvelle visite de l'enquêteur au ménage, afin d'obtenir des informations exactes⁵. Deuxièmement, les membres de l'équipe centrale de l'enquête devraient rendre des visites surprises aux équipes de terrain, afin de superviser les superviseurs, dont le travail a également besoin d'être contrôlé.

3. Gestion des données

29. L'une des tâches essentielles de toute enquête est d'entrer les données et de les présenter sous une forme qui en facilite l'analyse. Aujourd'hui, l'entrée des données se fait le plus souvent à l'aide d'ordinateurs personnels équipés d'un logiciel d'entrée des données. Ce logiciel doit être conçu pour éviter les incohérences dans les données. Si l'on constate des disparités, on peut au moins vérifier le travail des agents chargés d'entrer les données afin de déterminer s'il s'agit simplement d'erreurs de leur part. Le recours à un système encore plus efficace, prévoyant le retour de l'enquêteur auprès du

⁵ L'utilisation d'ordinateurs portables sur le terrain n'est pas nécessairement chose facile. On peut se heurter à des problèmes d'électricité, à des problèmes d'ordinateur causés par la poussière, la chaleur ou la forte humidité et, bien entendu, l'achat d'un grand nombre de ces ordinateurs est très coûteux.

ménage pour corriger ces disparités, serait possible si les données ont été entrées sur le terrain mais presque impossible si elles ont été entrées au siège central de l'organisation effectuant l'enquête.

30. Le système de gestion des données doit fonctionner de telle manière que les données parviennent le plus tôt possible à un centre. Cela est important pour deux raisons. Premièrement, il faudrait vérifier immédiatement le travail effectué la première semaine du premier mois pour s'assurer que les données qui parviennent au siège ne posent pas de sérieux problèmes. Deuxièmement, dans la plupart des cas, plus les informations parviennent rapidement aux analystes, plus elles ont de valeur.

31. On peut également donner quelques conseils plus spécifiques concernant la gestion des données. Premièrement, il faudrait tenir une comptabilité complète de tous les ménages de l'échantillon : de ceux qui ont répondu, de ceux qui n'ont pas répondu ou des unités jugées irrecevables. On a besoin de ces informations pour affecter une pondération aux données de ceux qui ont répondu pour l'analyse. Deuxièmement, le logiciel d'entrée des données devrait être soigneusement testé avant usage. L'excellent moment pour le faire serait lors de l'essai pilote du questionnaire. Troisièmement, avant de livrer l'ensemble des données aux chercheurs et aux analystes, il conviendrait d'en vérifier chaque partie pour s'assurer qu'aucun ménage n'a été exclu par erreur, ou inclus plus d'une fois. Quatrièmement, il y a lieu de préparer un document contenant des « informations de base » et de le remettre aux analystes, afin de s'assurer qu'ils comprennent comment utiliser les données. Cela est expliqué de façon plus détaillée à la section D.

D. Activités nécessaires une fois que le travail de terrain, l'entrée et le traitement des données sont terminés

32. Une fois que toutes les entrevues ont eu lieu, il y a encore certaines choses à faire pour assurer le succès d'une enquête sur les ménages. Toutes ces tâches se déroulent généralement au siège de l'organisation qui a recueilli les données. La tâche la plus évidente est l'analyse des données, qui fait l'objet d'un examen détaillé dans un autre chapitre de cette publication, mais plusieurs autres tâches finales doivent également être exécutées.

1. Compte rendu oral de mission

33. Tous les superviseurs et, si possible, tous les enquêteurs et agents d'entrée des données devraient participer à une réunion avec l'équipe centrale de l'enquête pour discuter de problèmes qui se sont posés, échanger des idées sur la possibilité de les éliminer dans les enquêtes futures et, d'une façon plus générale, offrir des suggestions sur les moyens d'améliorer le travail d'enquête. Cette réunion devrait avoir lieu immédiatement après que l'enquête a été terminée et avant que les agents de terrain et d'entrée des données n'oublient les détails de leurs expériences. Un compte rendu détaillé devrait consigner les recommandations, afin qu'il puisse en être tenu compte lorsque la prochaine enquête de ce type sera organisée.

2. Préparation de l'ensemble des données et de la documentation finale

34. Les données de presque n'importe quelle enquête sur les ménages ont des chances d'être utiles pendant de nombreuses années, et l'organisation qui les a re-

cueillies comme d'autres organismes de recherche (ou des chercheurs individuels) pourraient très bien s'en servir pour rédiger de nombreux rapports ou analyses dans les années à venir. Pour éviter toute confusion, une version finale « officielle » de ces données devrait être établie et servir de base à toute analyse par tout organisme ou individu qui désirerait les utiliser. L'idéal serait que cette version finale soit prête de deux à trois mois après la collecte des données. Par conséquent, les données recueillies sur le terrain doivent être rigoureusement vérifiées et analysées afin d'y déceler toute erreur et anomalie qui peut avoir à être corrigée ou, du moins, signalée. Bien entendu, certaines erreurs pourraient n'être découvertes qu'au bout de plusieurs mois ou même de plusieurs années, auquel cas un ensemble de données révisées pourrait être établi pour toute analyse ultérieure.

35. Les analystes peuvent avoir à poser de nombreuses questions au sujet des données. Ces questions peuvent être secondaires, sur la façon dont les fichiers de données ont été organisés, ou beaucoup plus importantes, sur la façon dont les données ont été recueillies. Pour éviter d'être submergés de demandes d'éclaircissements qui pourraient prendre une bonne partie du temps de leur personnel, les organisations qui collectent les données devraient préparer un document expliquant comment celles-ci ont été collectées et comment les fichiers de données ont été organisés et formatés. Ce document contiendrait des descriptions des codes qui ne figurent pas sur les questionnaires de l'enquête ainsi que des explications des cas où la collecte des données s'est écartée des plans initiaux. En principe, ce document devrait indiquer en quoi l'échantillon final diffère de l'échantillon envisagé; autrement dit, combien de ménages n'ont pu être trouvés ou ont refusé de participer et (s'il y a lieu) combien de nouveaux ménages ont été choisis pour remplacer ceux qui n'ont pas été interrogés. En plus de ce document, le « jeu » type d'informations à l'intention de tout analyste de données devrait comprendre un exemplaire du questionnaire et de tous les manuels de formation.

36. La dernière question concernant la documentation pour de nombreux pays a trait à sa traduction dans plusieurs langues. Aujourd'hui, de nombreux chercheurs étudient des pays dont ils ne lisent pas la langue, utilisant des traductions des questionnaires et autres documents. Au lieu d'attendre des chercheurs qu'ils fassent leur propre traduction, qui risque d'être inexacte, il est généralement préférable de traduire tous les documents nécessaires à l'analyse des données dans une langue internationale commune, le choix le plus évident étant l'anglais (mais le français et l'espagnol peuvent également être retenus). Ce travail peut s'avérer coûteux, mais on peut envisager d'inclure le coût de cette traduction dans le budget initial de l'enquête et demander aux donateurs qu'ils fournissent des fonds spécialement à cet effet.

3. Analyse des données

37. Toutes les données étant recueillies pour être analysées, il n'est pas nécessaire de préciser que la dernière activité après leur collecte est leur analyse. Comme cette question est traitée dans de nombreux autres chapitres, nous nous abstenons de le faire ici. Nous nous bornerons à souligner que le plan d'ensemble de l'enquête doit donner une estimation réaliste du temps nécessaire à l'analyse des données et incorporer ce temps dans le calendrier des activités de l'enquête. L'analyse des données prend presque toujours plus longtemps que prévu, mais plus les consultations sont étroites entre l'équipe de l'enquête et les personnes qui analysent ces données, plus les conclusions fondées sur ces données ont de chances d'être précises, et donc utiles.

E. Conclusion

38. Ce chapitre a offert des recommandations générales sur l'exécution d'enquêtes sur les ménages dans les pays en développement. La discussion a porté sur de nombreux sujets, mais chacun n'a été traité que brièvement – ce qui est inévitable, tant les enquêtes sur les ménages sont complexes. Comme les informations contenues dans ce chapitre sont insuffisantes pour une exécution minutieuse d'une enquête sur les ménages, quiconque envisage une telle enquête a besoin de consulter d'autres documents pour obtenir des conseils plus étayés. Il lui faut lire les ouvrages de référence cités dans l'introduction de ce chapitre; d'ailleurs, il est toujours judicieux de discuter de l'expérience d'enquêtes passées effectuées dans le pays en question avec les personnes ou les groupes qui ont effectué ces enquêtes. L'exécution d'une enquête peut être une tâche laborieuse, mais un travail appliqué, l'attention portée aux détails et l'application des conseils donnés dans le présent chapitre peuvent être déterminants pour la qualité et, par conséquent, l'utilité des données recueillies.

RÉFÉRENCES

- Casley, Dennis, et Denis Lury (1987). *Data Collection in Developing Countries*. Oxford, Royaume-Uni, Clarendon Press.
- Cochran, William (1977). *Sampling Techniques*. 3^e éd. New York, Wiley.
- Grosh, Margaret, et Juan Muñoz (1996). *A Manual for Planning and Implementing the Living Standards Measurement Study Survey*. Living Standards Measurement Study Working Paper, n° 126. Washington, D.C., Banque mondiale.
- Kalton, Graham (1983). *Introduction to Survey Sampling*. Beverly Hills, Californie, Sage Publications.
- Kish, Leslie (1965). *Survey Sampling*. New York, Wiley.
- Lohr, Sharon (1999). *Sampling: Design and Analysis*. Pacific Grove, Californie, Duxbury Press.
- Nations Unies (1984). Handbook of Household Surveys (revised edition). *Studies in Methods*, n° 31. Numéro de vente : E.83.XVII.13.

Section B

Conception de l'échantillon

Introduction

VIJAY VERMA

Université de Sienne
Sienna, Italie

1. La section A de cette publication donnait une vue d'ensemble des principaux problèmes techniques liés à la conception et à l'exécution d'enquêtes sur les ménages. Outre la conception du questionnaire, elle donnait un aperçu général des questions concernant l'exécution de l'enquête et la conception de l'échantillon. Dans la présente section, nous examinerons de façon plus spécifique certaines questions liées à la conception d'échantillons pour les enquêtes sur les ménages dans les pays en développement et les pays en transition. Cette section comprend trois chapitres : un chapitre sur la conception du cadre directeur d'échantillonnage et d'échantillons-maîtres pour les enquêtes sur les ménages; et deux chapitres concernant l'estimation des composantes des effets de la conception et leur utilisation dans la conception des échantillons.

2. L'enquête par échantillonnage a pour but de présenter des estimations ou des inductions applicables à l'ensemble de la population étudiée, à partir d'observations faites sur un nombre limité d'unités de cette population (l'échantillon). Ce processus est sujet à plusieurs types d'erreurs découlant de diverses sources. On fait généralement une distinction entre les erreurs d'échantillonnage et les erreurs non liées à l'échantillonnage. Toutefois, du point de vue de l'ensemble du processus de l'enquête, on fait une distinction plus fondamentale entre les « erreurs de mesure » et les « erreurs d'estimation ». Les erreurs de mesure, qui se produisent lorsque ce qui est mesuré sur les unités incluses dans l'enquête s'écarte des valeurs réelles (vraies) correspondant à ces unités, mettent en cause la précision de la mesure au niveau des unités individuelles recensées dans le cadre de l'enquête, et sont axées sur la substance de l'enquête. Elles se distinguent des erreurs d'estimation qui découlent du processus d'extrapolation à partir des unités particulières recensées à l'ensemble de la population étudiée pour laquelle des estimations ou des inductions sont nécessaires. Les erreurs d'estimation, qui concernent la possibilité de généralisation des unités observées à la population visée, sont liées au processus de conception et de mise en œuvre de l'échantillon. Ces erreurs comprennent, outre la variabilité de l'échantillonnage, diverses distorsions liées à la sélection de l'échantillon et à la mise en œuvre de l'enquête, notamment aux erreurs liées à la couverture et à la non-réponse. Toutes ces erreurs sont au centre des préoccupations du statisticien par échantillonnage. Souvent, plusieurs enquêtes ou cycles d'enquête utilisent un cadre d'échantillonnage, un échantillon-maître, une conception d'échantillon et parfois même un échantillon d'unités communs. En pareils cas, les erreurs liées au processus d'échantillonnage tendent à être communes à ces enquêtes et moins tributaires du thème traité.

3. C'est cette distinction entre la mesure et l'estimation qui sert de base au choix des questions traitées dans cette section. Les chapitres de la section B portent sur

deux aspects importants de l'estimation : le cadre d'échantillonnage, qui détermine la qualité de la couverture de la population étudiée et influe sur le coût et l'efficacité des modèles d'échantillon qui peuvent être conçus; et l'effet de la conception, qui donne une mesure quantitative de cette efficacité et peut aider à établir un lien entre la structure de l'échantillon et les coûts de l'enquête. Il y a certes d'autres aspects de la conception, et il serait donc utile d'étudier les chapitres de cette section en se référant au cadre élaboré dans la section précédente, en particulier aux principes fondamentaux et aux méthodes de conception d'échantillons présentés au chapitre II.

4. Le chapitre V traite avec force détails pratiques des concepts d'échantillon-maître et de cadre directeur d'échantillonnage. La définition de la population à laquelle les résultats obtenus avec l'échantillon doivent être généralisés est un aspect essentiel de la planification et de la conception de l'enquête. La population à étudier doit donc être représentée sous une forme physique à partir de laquelle on puisse sélectionner des échantillons du type requis. C'est cette représentation qui constitue le cadre d'échantillonnage. Dans le cas le plus simple, ce cadre est simplement une liste concrète de toutes les unités de la population; dans les modèles plus complexes, la représentation du cadre peut être partiellement implicite, tout en étant elle aussi constituée de toutes les unités. Dans la pratique, le cadre requis est défini par rapport à la structure requise des échantillons et à la procédure de sélection de ceux-ci. Dans les cadres à phases multiples qui, dans le cas des enquêtes sur les ménages, sont essentiellement fondés sur les zones couvertes, la durabilité du cadre diminue à mesure que l'on descend dans la hiérarchie des unités. À un bout, le cadre d'échantillonnage primaire représente un gros investissement pour un usage durable. À l'autre bout, les listes d'unités ultimes (telles qu'adresses, ménages et, surtout, personnes) nécessitent de fréquentes mises à jour.

5. Le cadre pour la première phase de l'échantillonnage (appelé cadre d'échantillonnage primaire) doit couvrir l'ensemble de la population des unités primaires d'échantillonnage (UPE). Après la première phase de sélection, la liste d'unités à tout stade de rang inférieur n'est nécessaire que parmi les unités de rang supérieur choisies à la phase précédente. Pour des raisons d'économies et de facilité, plusieurs phases de cette tâche peuvent être combinées ou partagées entre un certain nombre d'enquêtes. L'échantillon résultant des phases ainsi partagées est appelé échantillon-maître. L'objectif est d'offrir un échantillon commun d'unités jusqu'à un certain stade à partir duquel un nouvel échantillonnage peut être effectué pour certaines enquêtes particulières. L'échantillon-maître a notamment pour but de permettre :

- a) D'économiser, en partageant les coûts d'élaboration et de maintien de cadres et de matériels d'échantillonnage entre différentes enquêtes;
- b) De réduire le coût de conception et de sélection des échantillons;
- c) De simplifier le processus technique de sélection d'échantillons individuels;
- d) De faciliter l'établissement de liens concrets et opérationnels entre différentes enquêtes, en particulier entre les cycles successifs d'une même enquête;
- e) De faciliter, ainsi que de restreindre et de maîtriser en tant que de besoin, la sélection d'échantillons multiples pour diverses enquêtes à partir du même cadre.

6. Il importe de souligner que, dans la pratique, les échantillons-maîtres ont leurs limites :

- a) L'économie risque d'être modeste lorsque la notion d'échantillon-maître ne peut être étendue aux stades inférieurs de l'échantillonnage, où les unités sont moins stables et les listes ou cadres correspondants ont besoin de fréquentes mises à jour;
- b) On ne peut réaliser une économie raisonnable que si l'échantillon-maître est utilisé pour plus d'une et, de préférence, pour de nombreuses enquêtes;
- c) L'utilisation efficace d'un échantillon-maître nécessite une planification à long terme, ce qui n'est pas facile à réaliser dans les pays en développement;
- d) Le manque de souplesse dans la conception d'enquêtes auxquelles puisse être appliqué un échantillon-maître commun peut être un problème;
- e) La complexité technique de conception d'échantillons individuels peut être de plus en plus grande; quoi qu'il en soit, la documentation relative à l'échantillon-maître doit être tenue de façon détaillée et précise.

7. Il est également possible d'étendre la notion d'échantillon-maître non plus à un simple échantillon mais à la population tout entière, à l'ensemble des UPE. Telle est la notion de cadre directeur d'échantillonnage examinée au chapitre V. L'investissement dans un tel cadre est justifié lorsque les cadres disponibles ne couvrent pas pleinement la population étudiée et/ou ne contiennent pas d'informations propres à permettre une sélection d'échantillons simple et efficace. L'utilisation d'un cadre directeur d'échantillonnage permet aussi d'atténuer les contraintes liées au type et à la taille des échantillons qui peuvent être sélectionnés à partir d'un échantillon-maître plus restreint.

8. Les chapitres VI et VII traitent de la notion importante d'effet lié à la conception. L'effet lié à la conception (ou sa racine carrée, qui est parfois dénommée le coefficient de conception) est une mesure sommaire d'ensemble de l'effet sur la variance d'une estimation de divers facteurs de complexité de la conception. Pour une statistique donnée, il se calcule comme le ratio de sa variance dans le cas de la conception effective à ce que cette variance aurait été avec un simple échantillon aléatoire (SEA) de la même taille. Il donne ainsi une mesure de l'efficacité de la conception. En prenant le ratio de la variance effective à celle d'un SEA, l'effet lié à la conception retire également l'effet de facteurs communs aux deux, tels que la taille de l'estimation et l'ampleur de la mesure, la variance de la population et la taille globale de l'échantillon. Cela rend la mesure plus « transposable » d'une situation (enquête, conception) à une autre. Ces deux caractéristiques de l'effet lié à la conception, en tant que mesure sommaire et mesure transposable de l'efficacité de la conception, contribuent à la grande utilité et à la large utilisation de cette mesure dans l'exécution pratique d'une enquête. Le calcul et l'analyse des effets liés à la conception, que ce soit pour nombre de statistiques ou pour les estimations de diverses sous-catégories de population, sont d'une valeur inestimable pour l'évaluation de la conception des échantillons actuels et de nouveaux échantillons.

9. Sans éliminer certaines sources importantes de variation de l'ampleur de l'erreur d'échantillonnage mentionnée précédemment, l'ampleur de l'effet lié à la conception reste tributaire d'autres caractéristiques de la conception, telles que le nombre et le mode de sélection des ménages ou des personnes dans les zones d'échantillonnage. Mais surtout, il ne faut pas perdre de vue que les effets liés à la conception sont propres à la variable ou à la statistique concernée. Aucun effet lié à la conception ne décrit à lui seul l'efficacité d'échantillonnage de la conception. Pour la même conception, différents types de variables et de statistiques peuvent avoir (et ont souvent) des valeurs très différentes d'effet lié à la conception, au même titre que différentes estimations

de la même variable sur différentes sous-catégories de population. Cette diversité des valeurs de l'effet lié à la conception d'une enquête à une autre comme au sein d'une même enquête est illustrée par l'éventail des résultats empiriques portant sur divers types de variables extraites de 10 enquêtes effectuées dans six pays, présentées au chapitre VII.

Chapitre V

Conception de cadres directeurs d'échantillonnage et d'échantillons-maîtres pour les enquêtes sur les ménages dans les pays en développement

HANS PETTERSSON

Statistics Sweden
Stockholm, Suède

RÉSUMÉ

Le présent chapitre traite des questions concernant la conception de cadres directeurs d'échantillonnage et d'échantillons-maîtres. L'introduction est suivie de plusieurs sections. La section B donne un bref rappel des raisons pour lesquelles on conçoit et utilise des cadres directeurs d'échantillonnage et des échantillons-maîtres; la section C contient un examen des principales questions liées à la conception d'un cadre directeur d'échantillonnage; et la section D porte sur les échantillons-maîtres et traite des décisions importantes à prendre au stade de la conception (choix des UPE, nombre des phases d'échantillonnage, stratification, répartition de l'échantillon entre les diverses couches, etc.).

Termes clés : cadre directeur d'échantillonnage, échantillon-maître, conception de l'échantillon, échantillon à niveaux multiples.

A. Introduction

1. Dans les pays en développement, les Offices nationaux de statistiques (ONS) sont généralement les principales sources de statistiques « officielles ». À ce titre, les ONS doivent prendre en compte un large éventail de besoins d'informations dans les domaines des statistiques démographiques, sociales et économiques. Les ONS utilisent différentes sources et différentes méthodes de collecte de données. Ils peuvent avoir accès dans une certaine mesure aux données et registres administratifs, mais les enquêtes par échantillonnage seront toujours un important moyen de collecte. La plupart des ONS des pays en développement effectuent chaque année plusieurs enquêtes. Certaines de celles-ci (par exemple, l'Étude sur la mesure des niveaux de vie, l'enquête démographique et sanitaire, l'Enquête en grappes à indicateurs multiples) sont de conception relativement normalisée, tandis que d'autres sont conçues « sur mesure » pour répondre à des demandes nationales spécifiques. Le besoin de planifier et de coordonner les opérations de l'enquête ont encouragé les efforts en vue d'intégrer les enquêtes dans des programmes d'enquêtes sur les ménages. Dans beaucoup d'ONS, la programmation ad hoc d'enquêtes a désormais laissé la place à des plans à long terme selon lesquels des enquêtes portant sur différents sujets ont lieu continuellement ou à intervalles réguliers. Le Programme de mise en place des dispositifs nationaux d'enquête sur les ménages de l'Organisation des Nations Unies [*National Household Survey Capability Programme* (NHSCP)] a joué un rôle important dans ce processus.

2. Le programme d'enquête sur les ménages permet d'intégrer de plusieurs manières la conception et le déroulement de l'enquête. On peut utiliser les mêmes concepts et définitions pour les variables apparaissant dans plusieurs enquêtes. On peut accroître l'efficacité d'utilisation du personnel et des installations en les partageant entre plusieurs enquêtes. L'intégration peut aussi aller jusqu'à utiliser des cadres d'échantillonnage et des échantillons communs pour toutes les enquêtes de programme. L'élaboration d'un cadre directeur d'échantillonnage [*master sampling frame* (MSF)] et d'un échantillon-maître [*master sample* (MS)] pour les enquêtes est souvent un élément important d'un programme intégré d'enquêtes sur les ménages.

3. L'utilisation d'un cadre directeur d'échantillonnage commun d'unités géographiques pendant la première phase de l'échantillonnage permet d'améliorer le rapport coût/efficacité du programme d'enquêtes sur les ménages. Le coût d'élaboration d'un bon cadre d'échantillonnage est généralement élevé; l'établissement d'un programme continu d'enquêtes permet aux ONS d'étaler les coûts de construction du cadre d'échantillonnage sur plusieurs enquêtes.

4. Le partage des coûts peut aller un pas plus loin si les enquêtes choisissent leurs échantillons comme sous-échantillons d'un échantillon-maître commun du MSF. L'utilisation d'un échantillon-maître pour la totalité ou la plupart des enquêtes permet d'alléger les coûts de sélection d'échantillons et de préparation de cadres d'échantillonnage à la deuxième phase et aux suivantes de sélection pour chaque enquête. Ces économies que permettent le MSF et le MS s'appliquent aussi aux enquêtes spéciales entreprises durant l'exécution du programme d'enquêtes et, en fait, également lorsque l'ONS n'a pas de programme formel d'enquêtes.

5. Le présent chapitre traite les questions concernant la conception de cadres directeurs d'échantillonnage et d'échantillons-maîtres pour les enquêtes sur les ménages. Le manuel des Nations Unies *Programme de mise en place des dispositifs nationaux d'enquête sur les ménages* [*National Household Survey Capability Programme: sampling frames and sample designs for integrated household survey programmes*] (Nations Unies, 1986) donne une bonne description des différentes phases du processus de conception, préparation et tenue d'un cadre directeur d'échantillonnage et d'un échantillon-maître. Ce manuel présente plusieurs études de cas dans une annexe. Il est recommandé au lecteur que cela intéresse de se reporter à cette publication pour un examen détaillé de cette question.

B. Cadres directeurs d'échantillonnage et échantillons-maîtres : aperçu général

1. Cadres directeurs d'échantillonnage

6. Comme il est indiqué au chapitre II, dans les pays en développement, la sélection des échantillons de ménages se déroule normalement en plusieurs phases. Les unités d'échantillonnage utilisées à la première phase sont appelées unités primaires d'échantillonnage (UPE). Il s'agit d'unités géographiques. Ce peuvent être des subdivisions administratives telles que districts ou quartiers, ou des zones délimitées dans un but précis comme les zones d'énumération (ZE) pour recensement. Les unités utilisées lors de la deuxième phase sont les unités secondaires d'échantillonnage (USE) choisies parmi les UPE sélectionnées. Les unités de la dernière phase d'un échantillonnage à phases multiples sont appelées unités ultimes d'échantillonnage (UUE). Il faut un cadre d'échantillonnage — liste des unités parmi lesquelles l'échantillon est sélectionné — pour chaque phase d'un échantillon à phases multiples. Le cadre d'échantillonnage des unités de la première phase doit couvrir l'ensemble de la population

étudiée, de façon exhaustive et sans redondances; en revanche, les cadres d'échantillonnage de la deuxième phase ne seraient nécessaires que pour les UPE sélectionnées lors de la phase précédente.

7. Si les UPE sont des unités administratives, une liste de ces unités existe peut-être ou pourrait en général être facilement constituée à partir des dossiers administratifs à utiliser comme cadre d'échantillonnage. Une telle liste ad hoc d'UPE pourrait être établie chaque fois qu'un échantillon est nécessaire. Toutefois, lorsqu'il doit y avoir une série d'enquêtes à effectuer sur une période donnée, il serait préférable de préparer et de tenir un cadre directeur d'échantillonnage à utiliser dans chaque cas. L'économie pourrait être considérable par comparaison avec la préparation ad hoc de cadres d'échantillonnage pour chaque occasion. En outre, comme le cadre sera utilisé pour plusieurs enquêtes, il sera plus facile de justifier le coût de son établissement et de son maintien et de consacrer des ressources à en améliorer la qualité.

8. Un cadre directeur d'échantillonnage (MSF) est essentiellement une liste d'unités géographiques couvrant la totalité du pays. Pour chaque unité, on peut disposer d'informations sur sa classification (urbaine ou rurale), l'unité de rang supérieur (par exemple, le district ou la province) à laquelle elle appartient, son chiffre de population et, éventuellement, sur d'autres caractéristiques. Il faut aussi disposer d'informations sur ses limites géographiques. Ainsi, pour les enquêtes sur les ménages effectuées en République démocratique populaire lao, le MSF contient une liste de quelque 11 000 villages. Pour chaque village, on dispose d'informations sur le nombre de ménages, le nombre d'hommes et de femmes, sur sa classification comme urbain ou rural (les subdivisions administratives des zones urbaines sont également appelées villages) et sur le district et la province auxquels il appartient. Le MSF indique également si le village est accessible par la route.

9. Le type le plus courant de MSF est celui où les zones d'énumération (ZE) sont les unités de base. En général, on dispose d'informations pour chaque unité qui sert de lien avec les unités de rang supérieur (subdivisions administratives). Un tel MSF permet de sélectionner directement des échantillons de ZE. Il permet également de sélectionner des échantillons de subdivisions administratives, puis des ZE au sein des subdivisions sélectionnées.

10. Un MSF à jour et souple présente d'autres avantages que les avantages de coût et de qualité évoqués précédemment. Il facilite la sélection rapide et aisée d'échantillons pour des enquêtes de divers types et il permet de répondre à diverses exigences de l'enquête concernant les échantillons. L'autre avantage est qu'un MSF bien tenu sera utile pour le prochain recensement de la population. Le recensement lui-même a besoin d'un cadre semblable à celui qui sera utilisé pour les enquêtes sur les ménages. La tâche de l'élaboration du cadre pour le recensement devrait être considérablement allégée si un cadre directeur d'échantillonnage a été utilisé dans l'intervalle entre deux recensements. La situation idéale est celle où le nouveau MSF est envisagé et élaboré pendant cette période, puis entièrement mis à jour lors du prochain recensement.

2. Échantillons-maîtres

11. À partir d'un cadre directeur d'échantillonnage, on peut choisir en toute indépendance les échantillons à utiliser pour différentes enquêtes. Toutefois, dans bien des cas, il est très avantageux de choisir un vaste échantillon ou échantillon-maître, puis de sélectionner des sous-échantillons à partir de celui-ci pour des enquêtes différentes (mais apparentées). De nombreux ONS ont décidé d'élaborer un échantillon-maître pour les besoins de leurs enquêtes sur les ménages.

12. Un échantillon-maître est un échantillon à partir duquel on peut choisir des sous-échantillons pour les besoins de plus d'une enquête ou d'un cycle d'enquêtes (Nations Unies, 1986), et il peut prendre plusieurs formes. Un échantillon-maître de conception simple et assez courante est composé d'UPE qui sont des ZE. L'échantillon est utilisé pour un processus de sélection en deux phases, où les unités d'échantillonnage de la seconde phase (USE) sont des unités d'habitation ou des ménages.

13. La sélection de sous-échantillons peut se faire de différentes manières. Au niveau des UPE, à partir des UPE de l'échantillon-maître, elle produirait un sous-échantillon unique pour chaque enquête, de sorte que chacune de celles-ci disposerait d'un échantillon de différentes ZE. La sélection de sous-échantillons de niveau secondaire produirait un sous-échantillon d'unités d'habitation à partir de chaque UPE de l'échantillon-maître; autrement dit, chaque enquête aurait le même échantillon de ZE mais des échantillons différents d'unités d'habitation au sein de chaque ZE. Cette sélection pourrait se faire en toute indépendance, mais on pourrait aussi recourir à un processus de sélection contrôlée de façon à s'assurer que les chevauchements entre échantillons restent au niveau désiré. Un autre moyen de sélectionner des échantillons à partir de l'échantillon-maître serait de choisir des copies indépendantes de l'échantillon. On pourrait choisir une ou plusieurs copies comme sous-échantillon pour chaque enquête. Une telle formule exigerait que l'échantillon-maître soit constitué dès le départ à partir d'un ensemble de copies parfaitement indépendantes.

14. L'ONS peut réaliser des économies substantielles en utilisant un échantillon-maître. Les coûts de sélection d'unités de l'échantillon-maître seraient partagés entre toutes les enquêtes utilisant celui-ci; les coûts de sélection d'échantillons seraient donc plus faibles pour chaque enquête. Comme la sélection d'unités de l'échantillon-maître est essentiellement un travail de bureau (surtout si l'on dispose d'un bon MSF), l'économie réalisée à ce stade serait probablement modeste. En revanche, les économies sont beaucoup plus substantielles lorsque les coûts d'établissement de cartes et de cadres de sous-échantillonnage d'unités d'habitation au sein des unités de l'échantillon-maître sont partagés entre les enquêtes. L'établissement des cadres de sous-échantillonnage nécessite généralement un travail de terrain considérable, et le coût de ce travail par enquête diminue à mesure qu'augmente le nombre d'enquêtes utilisant le même cadre de sous-échantillonnage.

15. Dans certains pays, en raison des difficultés et des coûts que comportent les déplacements sur le terrain, il pourrait être plus économique de recruter les enquêteurs au sein ou à proximité des unités primaires d'échantillonnage de l'échantillon-maître et de les laisser sur place pour toute la durée de l'enquête. Dans ce cas, on utilise des UPE relativement grandes, et il est nettement préférable d'avoir un échantillon-maître fixe de ces UPE plutôt que de choisir un nouvel échantillon pour chaque enquête et d'avoir à localiser les enquêteurs ou à en recruter de nouveaux.

16. L'utilisation des mêmes unités de l'échantillon-maître permet de réduire le temps nécessaire au démarrage de l'enquête dans une zone donnée. Dans beaucoup de pays en développement, l'enquêteur doit obtenir l'autorisation des autorités régionales et locales de procéder à des entrevues dans la zone en question. Dans des pays comme la République démocratique populaire lao et le Viet Nam, par exemple, il lui faut obtenir des autorisations à plusieurs niveaux administratifs en descendant jusqu'au chef de village. Le délai nécessaire à ce processus « d'installation » diminue sensiblement lorsque les mêmes zones sont utilisées pour plusieurs enquêtes.

17. L'utilisation des mêmes UPE de l'échantillon-maître permet de réduire le temps nécessaire à l'enquêteur pour trouver les ménages. Lorsque l'on dispose de cartes et de cadres de sous-échantillonnage de bonne qualité, l'enquête peut facilement

trouver son chemin dans la zone d'enquête; dans certains cas, il ou elle peut même avoir déjà travaillé dans cette zone lors d'une enquête précédente. Une numérotation permanente des unités d'habitation peut être introduite pour faciliter son orientation. Cela a déjà été fait dans certains échantillons-maîtres : Torene et Torene (1987) décrivent le cas de l'échantillon-maître au Bangladesh.

18. L'échantillon-maître permet d'avoir des chevauchements d'échantillons dans deux enquêtes ou plus. Cela permet d'intégrer les données au niveau « micro » en liant les données sur les ménages provenant de ces enquêtes. Toutefois, lorsque des unités d'échantillonnage sont utilisées plusieurs fois, cela risque de nuire à la qualité des résultats de l'enquête. Les ménages participant à plusieurs cycles d'une enquête ou à plusieurs enquêtes peuvent se montrer réticents à participer ou moins enclins à donner des réponses exactes lors des enquêtes ultérieures.

19. L'échantillon-maître présente donc des avantages (coûts, intégration et coordination) pour les enquêtes ordinaires d'un programme donné. Il permet à l'ONS d'être mieux préparé à effectuer un échantillonnage pour des enquêtes spéciales : les sous-échantillons peuvent être choisis rapidement.

20. L'échantillon-maître présente des avantages visibles mais il présente aussi des inconvénients et des limites. Sa conception représente toujours un compromis entre différentes exigences découlant des enquêtes du programme. Il convient pour des enquêtes qui présentent des exigences de conception raisonnablement compatibles en ce qui concerne les estimations du domaine et la répartition de la population visée à l'intérieur de ces zones. La conception choisie pour l'échantillon-maître se prête généralement bien à la plupart des enquêtes du programme, mais jamais parfaitement. La conception de l'échantillon-maître impose aux différentes enquêtes des contraintes et des exigences (concernant la taille de l'échantillon, la mise en grappes, la stratification, etc.) auxquelles il est parfois difficile de répondre. Il en résulte une perte d'efficacité au niveau des enquêtes.

21. Certaines enquêtes présentent en outre certaines exigences spéciales de conception auxquelles l'échantillon-maître ne pourra pas répondre, à savoir :

- Les enquêtes portant sur certaines zones régionales ou locales nécessitant un vaste échantillon pour une petite zone (par exemple, les enquêtes utilisées pour évaluer les effets d'un projet de développement sur une zone locale donnée);
- Les enquêtes portant sur une population présentant des disparités (par exemple, ethniques).

22. Comme exemple du premier type, on peut citer l'enquête sur la culture du pavot, destiné à la production d'opium, effectuée régulièrement dans certaines zones de quatre provinces de la République démocratique populaire lao. Il s'agit de mesurer les progrès réalisés par le projet du Gouvernement lao de réduction de cette culture. Dans ce cas, comme l'échantillon-maître lao ne pouvait répondre aux exigences concernant la conception de l'échantillon, un échantillon différent a été sélectionné pour l'enquête. (On aurait pu également utiliser des UPE de l'échantillon-maître pour les quatre provinces et choisir d'autres UPE à partir du cadre de cet échantillon.)

23. Dans certains cas, il peut arriver que l'on ne réalise pas pleinement l'économie que permet l'échantillon-maître. Il faut des compétences techniques pour prélever un sous-échantillon à partir d'un échantillon-maître afin de répondre aux besoins particuliers d'une enquête donnée, puis calculer correctement les probabilités de sélection. Cette tâche peut se révéler plus compliquée que de choisir un échantillon indépendant. Le fait que beaucoup d'ONS de pays en développement ont peu de statisticiens par échantillonnage peut freiner l'utilisation, voire l'établissement d'un échan-

tillon-maître. Il existe des exemples d'échantillons-maîtres qui sont sous-utilisés en raison du manque de compétences en échantillonnage au sein des ONS.

3. Résumé et conclusion

24. Les avantages, inconvénients et limitations examinés ci-dessus peuvent se résumer de la façon suivante :

Cadre directeur d'échantillonnage :

- Économique; permet à l'ONS d'étaler les coûts de construction d'un cadre d'échantillonnage sur plusieurs enquêtes;
- La qualité est généralement meilleure que celle de cadres d'échantillonnage spéciaux car il est plus facile de mobiliser des investissements pour améliorer la qualité d'un cadre qui sera utilisé plus longtemps;
- Il simplifie le processus technique de sélection d'échantillons individuels et facilite une sélection rapide et aisée d'échantillons pour des enquêtes de types différents;
- S'il est bien tenu, il sera utile pour le prochain recensement de la population.

Échantillon-maître :

- *Économie :*
 - Les coûts de sélection d'unités de l'échantillon-maître seront partagés entre toutes les enquêtes utilisant cet échantillon.
 - Les coûts d'établissement de cartes et de cadres de sous-échantillonnage d'unités d'habitation ou de ménages seront partagés entre toutes les enquêtes utilisant l'échantillon-maître; cependant, les cadres de sous-échantillonnage devront être mis à jour périodiquement de manière à y ajouter les nouvelles constructions et à en retirer les unités d'habitation démolies.
 - Les gains tirés de l'utilisation d'un échantillon-maître sont clairs lorsque les enquêteurs doivent être placés à l'intérieur ou à proximité des UPE en raison des difficultés et du coût élevé des déplacements pour se rendre sur le terrain.
- *Efficacité accrue des opérations :*
 - L'utilisation des mêmes UPE de l'échantillon-maître pour plusieurs enquêtes réduit le temps nécessaire au démarrage des enquêtes dans la zone et le temps nécessaire à l'enquêteur pour trouver les ménages à interroger.
 - L'échantillon-maître permet un choix rapide et aisé d'échantillons; on peut choisir rapidement des sous-échantillons de l'échantillon-maître en cas de besoin pour des enquêtes spéciales.
- *Intégration :*
 - En permettant des chevauchements entre échantillons de deux enquêtes ou plus, l'échantillon-maître facilite l'intégration des données de ces enquêtes.

- *Limitations, inconvénients :*
 - L'échantillon-maître ne se prête pas à toutes les enquêtes; dans certains cas, pendant que se déroule le programme d'enquêtes, l'ONS se trouve dans des situations où apparaissent des besoins imprévus auxquels ne peut répondre l'échantillon-maître (cette limitation n'est pas vraiment un inconvénient).
 - Lorsque des unités d'échantillonnage sont réutilisées, notamment au niveau des ménages, il y a des risques de distorsions résultant d'effets de conditionnement et de l'accroissement du taux de non-réponse causé par le fait que les ménages sont sollicités à plusieurs reprises.
 - L'utilisation continue de l'échantillon-maître nécessite des compétences en échantillonnage dont ne dispose pas nécessairement l'ONS.

Conclusion

25. Il est clair que les cadres directeurs d'échantillonnage et les échantillons-maîtres présentent de nombreux attraits. Il est souhaitable que chaque ONS dispose d'un cadre directeur d'échantillonnage bien tenu qui puisse répondre aux besoins de ses enquêtes sur les ménages, que ces enquêtes fassent partie d'un programme ou soient effectuées à titre spécial. Nombre d'ONS jugeront de leur intérêt d'aller un pas plus loin et de concevoir et utiliser un échantillon-maître pour l'ensemble ou la plupart de leurs enquêtes sur les ménages.

C. Conception d'un cadre directeur d'échantillonnage

26. Le programme national d'enquêtes sur les ménages définit ce qui est attendu du cadre directeur d'échantillonnage et de la conception de l'échantillon-maître du point de vue, par exemple, du nombre prévu d'échantillons, de la population à couvrir, de la stratification et de la taille des échantillons. La façon dont il sera répondu à ces exigences dépend des conditions de construction du cadre dans le pays. La condition la plus importante est que l'on dispose de données et autres documents à utiliser pour la construction de ce cadre. À la section 1 ci-dessous, nous examinerons brièvement les types de données et de documents nécessaires et les problèmes de qualité que peuvent poser les données.

27. Lorsque les données et documents disponibles ont été évalués, l'ONS doit décider des principales caractéristiques du MSF concernant :

- La couverture du MSF (voir section 2);
- Quelles unités de zone devraient servir d'unités d'échantillonnage dans le MSF (voir section 3);
- Quelles informations concernant les unités du cadre devraient être incluses dans le MSF (voir section 4).

28. Une documentation complète et bien utilisée, ainsi que des procédures clairement définies pour sa mise à jour, sont cruciales pour la bonne utilisation du MSF (voir section 5).

1. Évaluation de la qualité des données et documents

29. La principale source de données et de documents est généralement le dernier recensement de la population. Cela est évident lorsque l'ONS a l'intention d'utiliser les zones d'énumération du recensement comme unités d'échantillonnage, mais

même si d'autres unités (administratives) sont utilisées, on a généralement besoin des données sur la population ou les ménages provenant du recensement. Les documents de base du recensement sont les listes de zones d'énumération (ZE). Il y a également les cartes des grandes zones (districts, régions) sur lesquelles sont marquées les ZE. Généralement, celles-ci sont identifiées par un code indiquant leur classification urbaine ou rurale ainsi que la division et la subdivision administrative à laquelle elles appartiennent. Parfois, le code indique également si la ZE contient des personnes vivant en caserne ou dans des foyers d'étudiants, etc.

30. La qualité des données et documents du recensement varie considérablement d'un pays à un autre. Cela est particulièrement le cas pour les cartes. Certains pays, comme l'Afrique du Sud, ont des cartes numérisées de leurs ZE conservées en mémoire sur des bases de données, tandis que d'autres, comme la République démocratique populaire lao, n'ont aucune bonne carte. Dans certains pays, les cartes des ZE sont souvent très schématiques et difficiles à utiliser sur le terrain. Comme les ZE peuvent en fait être composées de listes de localités plutôt que d'unités géographiques proprement dites, les populations éparses vivant à l'extérieur des localités figurant sur ces listes risquent de ne pas être couvertes par de tels cadres. La difficulté d'obtention d'informations provenant du recensement, en particulier de cartes, pose un problème particulier de qualité qui peut être gênant pour la conception du cadre. Les cartes peuvent être de bonne qualité mais cela n'aide pas davantage si elles sont difficiles à obtenir. Le fait qu'il est encore courant que les cartes des ZE soient « enfouies » dans des archives après le recensement, et parfois de façon désordonnée, les rend difficiles à trouver. Il n'est pas rare non plus que certaines cartes de ZE ne figurent pas dans les archives.

31. Généralement, la qualité des données du recensement se dégrade dans le temps. Tel est indiscutablement le cas des comptages des populations des ZE où la croissance démographique et les migrations influent de façon inégale sur ces populations. En outre, certaines modifications apportées à des unités administratives, telles que les changements de limites ou les fusions ou fractionnements d'unités, rendent caduques les données du recensement. Ces données ont toutes les chances de ne plus être valables si le dernier recensement remonte à sept ou huit ans.

32. La première étape dans la conception du MSF doit être d'identifier et d'évaluer les différents documents dont on dispose pour la construction du cadre, y compris non seulement les documents du recensement mais aussi d'autres données et documents : même si le recensement de la population doit être la principale source de documents, d'autres sources peuvent s'avérer nécessaires pour actualiser ou compléter les données du recensement. Les questions à poser sont : De quelles données et quels documents dispose-t-on et dans quelle mesure sont-ils exacts ? Ces données sont-elles d'actualité, et quelle est la fréquence de leur actualisation ? Les cartes ont besoin d'être évaluées pour déterminer leur degré de détail et dans quelle mesure les limites des subdivisions administratives y sont indiquées. Il faudrait s'efforcer d'estimer la proportion des cartes de ZE qui répondent aux normes de qualité requises.

33. À ce stade des travaux, il importe également d'obtenir ou de préparer une description précise et minutieuse de la structure administrative du pays et une liste à jour de ses divisions et subdivisions administratives.

2. Décision relative à la couverture du cadre directeur d'échantillonnage

34. Il faut décider très tôt de la couverture du MSF. Certaines parties reculées et à population éparses doivent-elles être exclues du cadre ? La décision de la plupart

des pays d'avoir une couverture nationale complète par le MSF est généralement avérée car, lorsque certaines régions reculées et à population éparses sont exclues des enquêtes ordinaires du programme, il peut encore être nécessaire de prévoir des enquêtes spéciales pour les couvrir. Les groupes nomades et les tribus de montagne difficiles à atteindre par les équipes de terrain sont un cas particulier. Dans certains pays, ces groupes sont exclus de la population visée par les programmes d'enquêtes sur les ménages.

35. Il faut également prendre une décision au sujet des populations vivant en institutions. Dans certains pays, les grandes institutions sont définies comme zones d'énumération spéciales (pensionnats, grands hôpitaux, casernes et foyers pour travailleurs des mines). Dans ce cas, on peut exclure ces zones du cadre. Toutefois, en général, il est préférable de conserver ces unités dans le cadre et de laisser ainsi de la place pour les décisions de couverture d'enquêtes à venir.

3. Décision concernant les unités de base du cadre d'échantillonnage

36. Les unités du cadre d'échantillonnage sont les unités figurant dans le cadre directeur d'échantillonnage. Les unités de base de ce cadre sont les unités les plus élémentaires du cadre directeur d'échantillonnage. Il est généralement souhaitable que ces unités soient de petites zones propres à permettre un groupement d'unités en unités d'échantillonnage plus grandes si une enquête particulière l'exige pour des questions de coût.

37. Les zones d'énumération du recensement sont souvent le meilleur choix d'unités de base du cadre d'échantillonnage. La délimitation de ces ZE a pour but de produire des ZE de tailles sensiblement égales du point de vue de leur population, ce qui présente un avantage dans certains cas d'échantillonnage. La carte des ZE est dressée, et généralement complétée par une description de leurs limites. Il existe généralement des cartes de base, indiquant l'emplacement des ZE au sein des divisions administratives. Il est établi des listes informatisées des ZE du recensement; ces listes peuvent être utilisées comme point de départ d'un MSF. Il y a beaucoup d'avantages à utiliser les ZE comme unités d'échantillonnage, mais des problèmes de qualité du type de ceux examinés à la section 1 peuvent parfois inciter à choisir d'autres solutions.

38. Certains pays ont des subdivisions administratives suffisamment petites pour pouvoir servir d'unités de base du cadre d'échantillonnage; et il peut y avoir des situations où ces unités offrent des avantages par rapport aux ZE comme unités de base de ce cadre; tel est le cas du MSF tenu par le Centre national de statistiques de la République démocratique populaire lao. Les ZE avaient été considérées comme unités de base du cadre d'échantillonnage, mais il a été constaté que la documentation des ZE était difficile à obtenir, et généralement de qualité médiocre, ce qui rendait les limites des ZE difficiles à déterminer sur le terrain. Il a donc été décidé d'utiliser les villages comme unités de base du cadre d'échantillonnage. Les villages de la République démocratique populaire lao sont des unités administratives bien définies. Toutefois, ils ne sont pas des unités géographiques au sens strict du terme. Les limites entre les villages sont floues, et il n'existe pas de cartes à proprement parler; néanmoins, il n'y a aucun doute sur le point de savoir à quel village appartient tel ou tel ménage.

39. Les cas où des unités plus petites que les ZE servent d'unités de base du cadre d'échantillonnage ne sont pas courants, mais ils existent. On peut citer l'exemple de la Thaïlande, où les ZE des zones municipales sont subdivisées en blocs et le recensement de la population et des ménages se fait pour chaque bloc. Ces blocs ont

été utilisés comme unités de base du cadre d'échantillonnage dans la partie municipale du MSF.

40. Les unités de base du cadre d'échantillonnage, qu'il s'agisse des ZE ou d'autres types d'unités, diffèrent par leur taille définie en nombre de ménages et de personnes de la zone. Même si l'on cherche à créer des ZE qui n'affichent pas de trop grosses différences de population, on ne peut manquer de déroger à cette règle et ce, pour diverses raisons (par exemple, des ZE moins peuplées peuvent être définies dans les régions à population éparses difficiles d'accès). Cela entraîne généralement de sensibles différences de taille entre les ZE, avec des cas extrêmes aux deux extrémités du spectre. Par exemple, au Viet Nam, le nombre moyen de ménages par zone d'énumération est de 100. Le nombre effectif de ménages de 166 000 ZE varie d'un minimum de 2 à un maximum de 304 (Glewwe et Yansaneh, 2001). Environ 1 % des ZE comptent 50 ménages ou moins. En République démocratique populaire lao, la proportion de petites ZE est encore plus grande : 6 % des ZE comptent moins de 25 ménages. Ces écarts de population entre les zones utilisées comme unités de base du cadre d'échantillonnage ne posent généralement pas de problème, mais les unités très petites ne conviennent pas comme unités d'échantillonnage. Les ZE très petites peuvent être acceptées dans le MSF, mais pour les échantillons fondés sur le MSF, ces ZE doivent être groupées avec des ZE adjacentes de manière à former des unités d'échantillonnage adéquates.

4. Informations sur les unités du cadre d'échantillonnage à inclure dans ce cadre

41. Une simple liste des unités de base du cadre d'échantillonnage constitue un cadre d'échantillonnage rudimentaire, mais la possibilité de prélever des échantillons efficaces de ce cadre est limitée. L'utilité de ce cadre serait beaucoup plus grande s'il contenait certaines données supplémentaires sur les unités du cadre qui pourraient servir à l'établissement d'échantillons efficaces. Ces données supplémentaires peuvent être de trois types :

- a) Informations qui permettent de grouper les unités de base du cadre d'échantillonnage en unités plus grandes. On peut accroître les possibilités d'échantillonnage efficace à partir du cadre en permettant l'échantillonnage de différents types d'unités du cadre. Il est donc souhaitable que le cadre contienne des informations qui permettent de former des unités plus grandes et d'assurer ainsi une certaine souplesse dans le choix des unités d'échantillonnage du cadre.
- b) Informations sur la taille des unités. On améliore aussi l'efficacité des échantillons du cadre si l'on indique la taille de chaque unité du cadre. Cela est particulièrement important lorsqu'il existe de grandes différences de taille entre ces unités.
- c) Autres informations supplémentaires. Les informations qui pourraient servir à la stratification des unités ou comme variables auxiliaires au stade de l'estimation permettent d'améliorer l'efficacité des échantillons du MSF.

Informations qui permettent de grouper les unités de base du cadre d'échantillonnage en unités plus grandes.

42. Pour certaines enquêtes, le meilleur choix pour les UPE est celui de petites zones comme zones d'énumération. Pour d'autres, les considérations de coût et d'erreurs d'échantillonnage militent en faveur d'UPE considérablement plus grandes que

les ZE. Ces UPE plus grandes pourraient être formées à partir de groupes de ZE voisines. On pourrait aussi utiliser comme UPE des unités administratives telles que quartiers et districts. Dans ces cas, il faut que le cadre directeur d'échantillonnage permette la construction de ces UPE plus grandes. Il est donc important que les registres des unités d'échantillonnage contiennent des informations sur les unités de niveau plus élevé auxquelles l'unité du cadre appartient.

43. L'un des modèles de cadre directeur d'échantillonnage qui a été utilisé par de nombreux pays est celui qui utilise les zones d'énumération du recensement comme unités de base du cadre d'échantillonnage et où ces unités sont ordonnées géographiquement en unités (administratives) plus grandes selon une structure hiérarchique. Les échantillons peuvent être extraits de différentes manières du MSF : a) par échantillonnage de ZE; b) en groupant les ZE de manière à former des UPE de taille convenable et en échantillonnant les UPE; et c) par échantillonnage des subdivisions administratives au premier stade puis à des stades supplémentaires en descendant jusqu'au niveau des ZE. La structure hiérarchique du cadre directeur d'échantillonnage du Viet Nam présente les niveaux suivants :

- Provinces
 - Districts
 - Communes (rurales), quartiers (urbains)
 - Villages (ruraux), blocs (urbains)
 - Zones d'énumération du recensement

44. La souplesse dans le choix des unités d'échantillonnage est encore accentuée si l'on affecte une identification à toutes les unités du cadre (les unités de base de même que les unités de niveaux plus élevés) sur la base de leur contiguïté géographique. Cela permet d'utiliser les unités du cadre comme modules pour la formation d'UPE de taille requise à partir d'unités du cadre adjacentes. Une telle opération serait nécessaire dans les cas du Viet Nam et de la République démocratique populaire lao décrits dans la section précédente. L'autre avantage de l'identification fondée sur la contiguïté géographique est que des échantillons géographiquement dispersés peuvent être choisis à partir du cadre directeur d'échantillonnage par l'utilisation d'un échantillonnage systématique à partir d'unités d'échantillonnage obéissant à un certain ordre géographique.

Mesures de la taille des unités du cadre d'échantillonnage

45. Le fait d'inclure la mesure de la taille est particulièrement important s'il y a de grandes différences de taille entre les unités. Généralement, cette mesure consiste en un comptage de la population, des ménages ou des unités d'habitation d'une unité du cadre. Il importe de souligner que la mesure de la taille n'a pas besoin d'être exacte. En fait, ces mesures sont presque toujours plus ou moins inexactes car elles reposent sur des données d'une époque antérieure et sont rendues peu à peu caduques par les variations constantes de la population. Les erreurs de mesure de la taille n'engendrent pas de distorsions dans les estimations de l'enquête mais elles réduisent l'efficacité de l'utilisation de ces mesures, surtout si on les utilise au stade de l'estimation. Il faudrait donc faire en sorte que ces mesures soient aussi précises que possible.

46. Les mesures de la taille des unités sont utilisées le plus couramment lors de la sélection d'unités du cadre selon une probabilité proportionnelle à leur taille. Elles sont également utilisées :

- Pour déterminer l'affectation des UPE de l'échantillon aux strates;

- Pour former des strates d'unités classées selon leur taille;
- Comme variables auxiliaires pour les estimations de ratio ou de régression;
- Pour former des unités d'échantillonnage d'une taille désirée.

Autres données supplémentaires pour les unités du cadre

47. On pourrait envisager d'inclure dans le cadre les informations supplémentaires sur des unités qui pourraient être obtenues moyennant un coût raisonnable. Les informations sur la densité de population, les groupes ethniques qui prédominent, la principale activité économique et le niveau moyen de revenu au sein des unités du cadre sont des variables qui sont souvent utiles pour la stratification.

48. Dans le cadre directeur d'échantillonnage de la Namibie, une classification brute selon les niveaux de revenu en revenu élevé, moyen et faible a été incluse pour les unités d'échantillonnage urbaines (ZE) de la capitale, Windhoek, ce qui a permis de former deux strates de niveau de revenu pour le sous-domaine urbain de Windhoek. L'autre exemple est celui du cadre directeur d'échantillonnage lao, dont les unités rurales contiennent des informations indiquant si l'unité est ou non proche d'une route. Les échantillons pour les enquêtes sur les ménages utilisant le cadre directeur d'échantillonnage sont stratifiés selon l'existence ou non d'un accès à une route.

5. Documentation et tenue d'un cadre directeur d'échantillonnage

Documentation

49. Une documentation bien tenue, précise et facile d'accès du cadre directeur d'échantillonnage est indispensable à l'utilisation de ce cadre. Une documentation médiocre ne permet pas de tirer pleinement profit des avantages du cadre. Au centre de cette documentation doit figurer une base de données dans laquelle sont incluses toutes les unités du cadre. Les informations sur ces unités devraient comprendre :

- Une identification de base, qui devrait être numérique, et avoir un code identifiant exclusivement toutes les divisions et subdivisions administratives dans lesquelles se situe l'unité. Il serait préférable que les unités soient numérotées selon un ordre géographique. Généralement, les codes des ZE présentent ces caractéristiques. Une identification entièrement numérique est préférable à un nom ou à un code alphanumérique. Souvent, les systèmes de codification géographique provenant de sources administratives et du recensement conviennent comme principale identification.
- Une identification secondaire, qui peut être le nom du village (ou de toute autre subdivision administrative) où est située l'unité. Cette identification secondaire est utilisée pour localiser l'unité sur les cartes et sur le terrain.
- Un certain nombre de caractéristiques de l'unité, telles que sa taille (population, nombre de ménages), son caractère urbain ou rural, sa densité de population, etc. Toutes les données concernant l'unité qui peuvent être obtenues moyennant un coût raisonnable et sont d'une qualité acceptable devraient être incluses. Ces caractéristiques pourraient être utilisées pour la stratification, pour l'affectation de probabilités de sélection, et comme variables auxiliaires dans l'estimation.
- Des données opérationnelles, des informations sur les changements survenus dans les unités et l'indication de leur utilisation dans l'échantillon.

50. Le cadre doit être facilement accessible et facile à utiliser pour diverses manipulations telles que le tri, le filtrage et la production de statistiques sommaires susceptibles de faciliter la conception et l'estimation de l'échantillon. Le meilleur moyen de lui donner ces caractéristiques est de le mettre en mémoire sur une base de données informatique. Il convient d'éviter d'utiliser des formats qui ne sont accessibles qu'aux spécialistes. Un simple tableur Excel est souvent suffisant. Excel est facile à utiliser, bien connu et offre les fonctions de tri, filtrage et agrégation nécessaires à la préparation d'échantillons à partir du cadre. Les tableurs peuvent facilement être importés par la plupart des autres logiciels.

Tenue du MSF

51. Il existe un lien étroit entre les opérations courantes nécessaires au maintien du cadre et la documentation du MSF. Pendant l'utilisation du MSF, il se produit des changements qui influent sur le nombre et la définition des unités du cadre. Le volume de travail nécessaire au maintien d'un cadre directeur d'échantillonnage dépend principalement de la stabilité de ces unités. Les unités du cadre peuvent subir deux types de changements : les modifications de leurs limites et les modifications de leurs caractéristiques.

52. Les modifications des limites des unités du cadre touchent principalement leurs subdivisions administratives. Celles-ci sont sujettes à des modifications de leurs limites, surtout aux niveaux inférieurs, par suite de décisions politiques ou administratives. Souvent, ces modifications font suite à des changements profonds de la population des zones touchées. De nouvelles unités sont créées par fractionnement ou combinaison ou par des réaménagements plus compliqués d'unités existantes. Par ailleurs, les limites d'unités existantes peuvent être modifiées sans que de nouvelles unités soient créées. S'il se produit de fréquents changements dans les subdivisions administratives, il faut allouer des ressources considérables pour tenir le cadre à jour et exact.

53. Les changements portant sur les rectifications des limites d'unités du cadre doivent être enregistrés dans le MSF. Il faut mettre en place un système de collecte d'informations pour suivre ces changements administratifs.

54. Les changements des caractéristiques des unités peuvent n'être que de simples changements, tels que changements de nom, ou être plus profonds et porter sur la taille (population ou nombre de ménages/unités d'habitation) ou sur la classification urbaine ou rurale. Ces changements ne doivent pas nécessairement être portés sur le MSF. Toutefois, comme on l'a vu précédemment, des informations caduques sur la taille des unités entraînent une baisse d'efficacité des échantillons sélectionnés à partir du cadre. L'actualisation des chiffres concernant la taille de toutes les unités du cadre serait très coûteuse et le plus souvent d'un faible rapport coût/efficacité, mais pour les zones périurbaines à croissance rapide, en particulier, il est bon d'actualiser régulièrement ces chiffres.

55. Les changements de taille des unités du cadre deviennent problématiques lorsqu'il se produit de vastes et soudaines variations de la population, comme ce peut être le cas, par exemple, dans les quartiers spontanés, lorsque les autorités locales décident de retirer les squatters de ces quartiers. De tels bouleversements doivent être signalés dans le cadre d'échantillonnage. La décision du Gouvernement de la République démocratique populaire lao de déplacer les populations des villages reculés des zones montagneuses est un exemple de changement moins frappant mais néanmoins problématique pour le cadre d'échantillonnage. Le gouvernement encourage les habitants de ces villages à s'installer dans des villages où les services de base sont plus ac-

cessibles. Sous l'effet de cette décision, le nombre de villages a diminué d'environ 10 % en deux ans. Il est clair que de tels changements doivent être indiqués dans le cadre d'échantillonnage.

56. Lorsque l'ONS dispose de ressources insuffisantes et se débat pour répondre à la demande de statistiques, il court le risque d'avoir à négliger la tenue à jour du MSF. Il importe donc qu'il dispose à un stade précoce de plans et de procédures d'actualisation de ce cadre et de ressources suffisantes à cet effet.

D. Conception d'échantillons-maîtres

57. L'échantillon-maître est un échantillon à partir duquel des sous-échantillons peuvent être sélectionnés pour répondre aux besoins de plusieurs enquêtes ou cycles d'enquêtes (Nations Unies, 1986). Le principal but doit être d'offrir des échantillons pour enquêtes sur les ménages présentant des caractéristiques de conception raisonnablement compatibles pour ce qui est des domaines d'analyse et de la distribution de la population visée à l'intérieur des zones sélectionnées. L'échantillon-maître se définit par le nombre de phases d'échantillonnage et le type d'unités devant servir d'unités ultimes d'échantillonnage (UUE). Un échantillon-maître sélectionné en deux phases, dont les zones d'énumération sont les unités de la seconde phase, est appelé *échantillon-maître de zones d'énumération à deux phases*. Si les ZE étaient sélectionnées directement durant la première phase, nous aurions un *échantillon-maître de zones d'énumération à une phase*. Ces deux types sont courants dans les pays en développement.

58. Les opérations importantes entrant dans l'élaboration d'un échantillon-maître sont examinées aux sections D.1-D.4. Les sections D.5 et D.6 traitent des questions concernant la documentation et la tenue de l'échantillon-maître. Enfin, la section D.7 est consacrée à l'utilisation de l'échantillon-maître pour les enquêtes qui ne s'adressent pas principalement aux ménages.

1. Choix des unités primaires d'échantillonnage à incorporer à l'échantillon-maître

59. Le MSF représente le cadre d'où est tiré l'échantillon-maître. Dans certains cas, l'unité de base du MSF pourrait servir d'unité primaire d'échantillonnage pour le choix de l'échantillon-maître. Dans d'autres cas, on pourrait décider de former des UPE plus grandes que les unités de base du MSF. Dans ces cas, on utilise certains types d'unités administratives bien définies (comtés, quartiers, etc.) comme UPE, mais il y a aussi les cas où l'on utilise les unités du cadre comme modules de construction. Dans ces cas, on groupe les unités adjacentes en UPE de taille appropriée. Ainsi, pour son échantillon-maître, le Lesotho a formé les UPE en combinant les ZE adjacentes du recensement en groupes de 300 à 400 ménages. Les 3 055 ZE du recensement ont été regroupées en 1 038 groupes de ZE qui devaient servir d'UPE (Pettersson, 2001).

60. Il faut tenir compte de plusieurs facteurs concernant l'efficacité statistique, les coûts et les procédures opérationnelles lorsque l'on décide de ce que devrait être l'unité primaire d'échantillonnage. Si l'on suppose que les unités primaires d'échantillonnage du MSF sont les ZE, dans quels cas préférerait-on utiliser des unités plus grandes que les ZE comme UPE ?

- Si nous savons que les limites d'une proportion importante de ZE ne sont pas bien définies, nous pouvons décider d'utiliser des unités plus grandes comme UPE car la délimitation des zones plus vastes est généralement plus stable et plus clairement établie;

- Lorsque les déplacements d'une zone à une autre sont difficiles et/ou coûteux. En raison des difficultés et des coûts des déplacements sur le terrain, il peut être plus économique de recruter les enquêteurs à l'intérieur ou à proximité des UPE sélectionnées pour l'échantillon et de les maintenir sur place pendant toute la durée de l'enquête. Cela supposerait le choix d'UPE assez vastes;
- Lorsque l'utilisation des UPE pour les échantillons serait si ample qu'une petite UPE telle qu'une ZE aurait tôt fait d'être épuisée. On peut résoudre ce problème soit en utilisant de plus grandes unités comme UPE ou en gardant les ZE comme UPE et en renouvelant l'échantillon de ZE. La première option est préférable lorsque le coût de lancement de l'enquête dans la zone est élevé;
- Lorsque, pour des raisons de limitation des coûts et d'efficacité d'échantillonnage, on utilise une ou plusieurs phases d'échantillonnage utilisant des unités plus grandes que les unités de base du cadre. Si, par exemple, les unités de base du cadre sont des ZE, on peut décider d'utiliser des unités plus grandes, notamment des quartiers, comme UPE, puis choisir des ZE ou d'autres unités géographiques au sein de l'UPE pour la phase suivante;
- Lorsque, comme dans certaines enquêtes, les variables du ménage ou des individus sont liées aux variables de la communauté. Tel est le cas des enquêtes sanitaires où les variables propres à la santé des individus sont liées aux variables concernant les équipements de santé du village ou de la commune. L'autre exemple est celui des enquêtes sur le niveau de vie où les variables concernant le ménage sont liées aux variables de la communauté en matière d'écoles, de routes, d'eau, d'assainissement, de prix locaux, etc. Si l'échantillon-maître sert à plusieurs enquêtes de ce type, il y a intérêt à utiliser la communauté (village, commune, quartier, etc.) comme UPE. Si tel est le cas, on peut veiller à ce que le sous-échantillon d'unités secondaires d'échantillonnage (USE) soit bien étalé sur l'ensemble de la communauté.

61. Les grandes unités ne peuvent convenir comme UPE car elles sont trop peu nombreuses. Il ne serait pas significatif de prélever un échantillon sur une population de 50 à 100 unités. Il serait préférable que le nombre d'UPE soit de plus de 1 000, de sorte qu'un échantillon de 10 % produirait plus de 100 UPE pour l'échantillon. Un pourcentage plus fort que 10 % réduirait l'avantage économique de l'échantillonnage. Un nombre d'UPE nettement inférieur à 100 dans l'échantillon augmenterait la variance. Il convient également de noter qu'il pourrait être efficace d'utiliser différents types d'UPE pour différentes catégories de population, par exemple les ZE pour les zones urbaines et de plus grandes unités dans les zones rurales.

2. Fusion/fractionnement des zones pour réduire les différences de taille entre UPE

62. Lorsque l'on a décidé du type d'unité qui devrait servir d'UPE (et, dans le cas où il y a deux catégories de zones, quelle unité devrait servir d'USE), on peut constater que certains « éléments atypiques » sont beaucoup plus petits ou beaucoup plus grands qu'il n'est souhaitable.

Unités d'échantillonnage très petites

63. L'utilisation d'UPE très petites dans l'échantillon-maître est problématique. Ce qui devrait être considéré comme une taille acceptable dépend de la charge de travail prévue pour l'échantillon-maître. Statistics South Africa, qui utilise les ZE du

recensement comme UPE pour son échantillon-maître a décidé de fixer la taille minimum des UPE à 100 ménages. Les ZE qui comptent moins de 100 ménages ont été rattachées à des ZE voisines pour la préparation du MSF. Pour son échantillon-maître, le National Central Statistics Office de la Namibie s'est fixé pour règle que les UPE devaient compter au moins 80 ménages. Lors du recensement, il a été formé 2 162 ZE. Après rattachement des petites ZE à des ZE voisines, il est resté 1 696 UPE. Sur ce total, 405 étaient le résultat de la fusion de plusieurs ZE, et les 1 291 UPE restantes étaient composées d'une seule ZE.

64. Le travail qu'implique le rattachement de petites ZE avant sélection peut être très laborieux s'il y a un grand nombre de petites ZE. À cet égard, on peut citer le cas du Viet Nam. Pour ses enquêtes, l'Office général des statistiques du Viet Nam voulait un échantillon de zones d'au moins 70-75 ménages. Environ 5 % des ZE (soit 8 000) comptent moins de 70 ménages (Pettersson, 2001b). Le travail de rattachement de quelque 8 000 ZE à des ZE voisines a été long et pénible.

65. Un moyen de réduire le travail de rattachement des petites ZE de manière à former des UPE de taille raisonnable est de n'effectuer cette opération que lorsqu'une petite zone (UPE) est choisie pour l'échantillon. Kish (1965) a conçu une procédure pour rattacher les petites UPE à des UPE voisines pendant ou après le processus de sélection.

66. Un autre moyen de réduire le travail de rattachement des petites ZE est d'introduire une phase d'échantillonnage avant la première phase prévue. Dans certains cas, au lieu d'utiliser les unités prévues comme UPE, on pourrait en utiliser de plus grandes. Dans les UPE sélectionnées, on combine de petites unités (prévues à l'origine comme UPE) en zones de taille raisonnable. Ce travail ne se fait que pour les unités sélectionnées lors de la première phase, ce qui réduit considérablement le travail, par comparaison au cas où on utilise les petites zones comme unités de première phase. Cette option implique une phase d'échantillonnage supplémentaire en amont de la première phase prévue, ce qui peut nuire à l'efficacité de la conception. Toutefois, si l'on ne choisit qu'une USE par UPE sélectionnée à la deuxième phase, l'échantillon est en fait équivalent à l'échantillon d'unités prévu avec une seule phase d'échantillonnage. C'est cette solution qui a été retenue dans le cas du Viet Nam. Il a été décidé d'utiliser des unités administratives plus grandes, en l'occurrence des communes, au lieu des ZE, comme UPE. Au sein des communes sélectionnées, les ZE de petite taille ont été rattachées à des ZE adjacentes pour former des unités de taille acceptable. Ainsi, le travail consistant à rattacher ces ZE a été allégé. Au lieu d'avoir à rattacher 8 000 ZE, le travail a été limité à environ 1 400 ZE des 1 800 communes sélectionnées. Trois ZE (ou groupes de ZE dans le cas de ZE de petite taille) ont été sélectionnées à la deuxième phase dans les communes sélectionnées.

Unités d'échantillonnage très grandes

67. À l'autre extrême, certaines unités peuvent être trop grandes – du point de vue de leur population ou de leur étendue géographique – pour servir d'UPE. Dans les deux cas, les coûts de recensement seront beaucoup plus élevés que pour des unités ordinaires (ZE ou autres). Des problèmes se posent dans un cas comme dans l'autre si certaines des très grandes UPE sont sélectionnées pour l'échantillon-maître. Pour alléger la tâche de préparation de liste des ménages de ces grandes unités, on peut mettre ces unités dans des strates séparées et les choisir moins fréquemment comme UPE; on peut maintenir les taux globaux d'échantillonnage en augmentant les taux d'échantillonnage à l'intérieur des UPE.

68. On peut également traiter le problème des grandes UPE en les divisant en un certain nombre de segments et en choisissant un segment au hasard. Le problème est légèrement plus simple qu'avec les petites UPE, car on n'a pas à prendre de mesures avant la sélection de l'échantillon-maître. On n'a de mesures à prendre que lorsque l'on choisit une grande UPE pour cet échantillon.

69. Les UPE qui ont grandi ou rétréci sensiblement depuis le dernier recensement posent un autre type de problème. Il se produit toujours des variations de population dans le temps, et cela réduit la précision de la mesure de la taille des UPE. La conséquence générale de cette situation est un accroissement des variances; en revanche, aucune distorsion n'est introduite. Le problème devient plus sérieux lorsqu'il se produit des changements spectaculaires dans certaines UPE, en raison, par exemple, de l'assainissement de certains quartiers suburbains ou de vastes programmes de construction de logements dans d'autres quartiers. Les procédures à employer pour tenir compte de ces variations doivent faire partie intégrante de la tenue de l'échantillon-maître. Le manuel du NHSCP propose deux stratégies : le remplacement de l'échantillon et la révision de l'échantillon (Nations Unies, 1986).

3. Stratification des UPE et répartition de l'échantillon-maître entre les strates

Stratification

70. Les UPE de l'échantillon-maître sont souvent stratifiées selon les principales divisions administratives du pays (provinces, régions, etc.) et au sein de ces divisions, en zones urbaines et rurales. La stratification se fait souvent également en fonction de certains facteurs tels que le niveau d'urbanisation (métropoles, grandes villes, petite villes et villages) et des caractéristiques socio-économiques et écologiques. Dans l'échantillon-maître du Lesotho, les UPE sont stratifiées en 10 régions administratives et 4 zones agroéconomiques (plaines, contreforts, montagnes et vallée du fleuve Senqu), ce qui a produit 24 strates reflétant les divers modes de vie dans les zones rurales.

71. On peut définir des strates correspondant aux « franges suburbaines » formées par les zones rurales à la périphérie des grandes villes. Cette catégorie correspond aux familles rurales qui sont plus ou moins tributaires du secteur moderne. Dans les grandes villes, on peut envisager une stratification secondaire établie selon la catégorie de logement, le niveau de revenu et certaines autres caractéristiques socio-économiques.

72. Une autre technique utilisée couramment pour établir une stratification plus détaillée à l'intérieur des principales strates est de classer les UPE d'une strate selon un critère de stratification donné et de choisir l'échantillon de façon systématique (stratification implicite). L'avantage de cette formule est qu'il n'est pas nécessaire de définir des limites des strates.

Allocation des UPE de l'échantillon-maître

73. L'allocation des UPE de l'échantillon-maître aux différentes strates peut prendre différentes formes :

- Allocation proportionnelle à la population de la strate;
- Allocation aux strates à parts égales;
- Allocation proportionnelle à la racine carrée de la population de la strate.

74. Dans de nombreux échantillons-maîtres, l'allocation aux strates se fait proportionnellement à la population (au nombre de personnes ou de ménages). L'allocation proportionnelle est souvent une bonne stratégie. Mais elle n'alloue qu'une faible proportion de l'échantillon aux petites strates. Cela peut poser un problème lorsque les principales strates sont des régions administratives (par exemple, des provinces) du pays pour lesquelles l'enquête nécessite des estimations séparées et lorsque ces régions diffèrent largement par leur taille (comme c'est souvent le cas). La demande d'allocation à égalité à l'échantillon entre toutes les provinces peut être très forte parmi les fonctionnaires des provinces (du moins, des petites provinces). Lorsque les provinces sont de tailles très différentes, l'allocation à égalité entraîne de substantiels écarts entre les fractions de l'échantillon représentant les diverses provinces. Pour l'échantillon-maître lao élaboré en 1997, il a été décidé d'utiliser l'allocation à égalité entre les 19 strates des provinces, de manière à obtenir la même précision dans les estimations établies pour chaque province. Cela s'est traduit par une situation où la province la plus petite avait une fraction d'échantillonnage 10 fois plus grande que la province la plus peuplée.

75. Une allocation strictement proportionnelle entre domaines urbains et ruraux produit de petits échantillons urbains dans les pays à faible population urbaine. L'échantillon-maître préparé par l'Institut national de la statistique du Cambodge est établi selon une allocation proportionnelle entre les provinces urbaines et rurales. L'échantillon de 600 UPE comprend 512 UPE rurales et 88 UPE urbaines. Pour certaines enquêtes, l'échantillon urbain a été jugé trop petit, et des UPE urbaines ont dû être ajoutées. Il pourrait être judicieux de gonfler quelque peu le domaine urbain dans l'échantillon-maître.

76. *L'allocation selon la racine carrée*, dans laquelle l'allocation se fait proportionnellement à la taille des strates, offre un compromis entre l'allocation proportionnelle et l'allocation à égalité. L'allocation selon la racine carrée a été utilisée pour les échantillons-maîtres au Viet Nam et en Afrique du Sud. Kish (1988) a proposé une autre solution de compromis fondée sur une allocation proportionnelle à $n\sqrt{(W_h^2 + H^{-2})}$ où n est la taille globale de l'échantillon, W_h est la taille relative de la strate h , et H est le nombre de strates. Pour chaque petite strate, le second terme domine le premier, faisant en sorte que les allocations aux petites strates ne soient pas trop petites.

77. On peut également envisager une solution de compromis où l'on aurait un grand échantillon-maître convenant aux estimations au niveau des provinces et un sous-échantillon, extrait de l'échantillon-maître, conçu principalement pour les estimations nationales. Tel est le cas de l'échantillon-maître conçu en 1996 aux Philippines, qui comprenait 3 416 UPE pour un échantillon élargi utilisé pour les estimations au niveau des provinces et un sous-échantillon de 2 247 UPE servant d'échantillon-maître de base pour les cas où l'on n'aurait besoin que d'estimations au niveau régional.

4. Sélection de l'échantillon d'UPE

78. La méthode la plus courante est de sélectionner les UPE de l'échantillon-maître selon une probabilité proportionnelle à leur taille. En l'occurrence, la probabilité de sélection d'une UPE est proportionnelle à la population de cette unité, ce qui donne aux grandes UPE une plus forte probabilité de figurer dans l'échantillon.

79. Cette méthode offre certains avantages pratiques lorsque les UPE sont de tailles très variables. Premièrement, elle permet une autopondération des échantillons. Deuxièmement, elle engendre des échantillons de tailles approximativement égales au sein des UPE; cela implique des charges de travail approximativement égales pour les

enquêteurs, ce qui est une situation souhaitable du point de vue du travail de terrain. Pour plus de renseignements sur l'échantillonnage selon une probabilité proportionnelle à la taille des UPE, se reporter au chapitre II.

80. L'échantillon fondé sur une probabilité proportionnelle à la taille des UPE peut être choisi de différentes manières. La sélection systématique à l'intérieur des strates est une méthode courante. Si les UPE sont classées selon un ordre géographique à l'intérieur des strates, on obtient une distribution géographique de l'échantillon à l'intérieur des principales strates (pour plus de détails, se reporter au chapitre II). Les échantillons-maîtres du Lesotho, de la République démocratique populaire lao et du Viet Nam ont tous été sélectionnés selon une probabilité proportionnelle à la taille des UPE, avec un point de départ choisi systématiquement au hasard dans chaque strate.

Interpénétration des sous-échantillons

81. Un autre moyen de former l'échantillon repose sur la sélection d'un ensemble de sous-échantillons à interpénétration. Un sous-échantillon à interpénétration est un sous-échantillon composé d'un ensemble de sous-échantillons dont chacun constitue en soi un échantillon de probabilité de la population visée.

82. La possibilité d'utiliser des sous-échantillons à interpénétration pour la sélection de sous-échantillons à partir de l'échantillon-maître présente certains avantages. Les sous-échantillons offrent une certaine souplesse concernant la taille de l'échantillon. L'échantillon choisi pour une enquête particulière peut être formé d'un sous-échantillon ou plus. Les sous-échantillons peuvent aussi être utilisés pour le remplacement de l'échantillon dans le cas d'enquêtes à cycles multiples.

83. L'utilisation de sous-échantillons à interpénétration pour la conception de l'échantillon-maître n'est pas aussi courante que l'utilisation d'une simple sélection systématique. L'échantillon-maître conçu par le Statistics Office du Nigéria offre un exemple d'échantillon-maître utilisant des échantillons à interpénétration.

5. Durabilité des échantillons-maîtres

84. La qualité de l'échantillon-maître se dégrade avec le temps, mais le fait que la mesure de la taille utilisée pour affecter les probabilités de sélection devienne caduque à mesure que se produisent des changements dans la population ne poserait pas de problèmes si ces changements étaient une croissance plus ou moins uniforme de toutes les unités du cadre directeur d'échantillonnage. Toutefois, tel est rarement le cas. La croissance de la population et les migrations ont des rythmes différents selon les zones : souvent, les zones rurales affichent une faible croissance, sinon un déclin, alors que les zones suburbaines de la périphérie des villes connaissent une croissance élevée. Lorsque la croissance est aussi inégale, les tailles utilisées lors de la sélection des unités de l'échantillon-maître cessent d'être représentatives de la répartition relative de la population étudiée. Cela entraîne un accroissement des erreurs d'échantillonnage dans les estimations tirées de l'échantillon-maître. En outre, les changements dans les limites et les classifications administratives (par exemple, zones urbaines ou rurales) peuvent faire perdre de son actualité à la stratification.

85. En principe, le cadre directeur d'échantillonnage est entièrement révisé après chaque recensement de la population, c'est-à-dire en général tous les dix ans. Entre les recensements, ce cadre devrait être régulièrement actualisé. Le fait de disposer d'un cadre directeur d'échantillonnage régulièrement actualisé permet de choisir périodiquement à partir de celui-ci des échantillons-maîtres entièrement nouveaux. Dès lors, la question qui se pose est de savoir combien de temps garder un échan-

tillon-maître sans lui apporter de changements significatifs ? La durabilité de l'échantillon-maître dépend, dans une certaine mesure, des conditions locales telles que les migrations internes et le rythme des changements qui se produisent à l'intérieur des unités administratives. On ne peut donc faire de recommandation générale applicable à toutes les situations. Souvent, au bout de trois à quatre ans, l'efficacité de l'échantillon-maître s'est sensiblement détériorée. Par conséquent, il faut réfléchir sérieusement avant de décider d'utiliser plus longtemps un échantillon-maître sans lui apporter d'ajustements.

86. Il y a essentiellement deux stratégies pour faire face au problème de la perte d'efficacité de l'échantillon-maître. La première est de choisir régulièrement un échantillon-maître entièrement nouveau; ainsi, au Lesotho, l'échantillon-maître est remplacé tous les trois ans. L'autre stratégie est de retenir l'échantillon-maître pendant plus longtemps mais de lui apporter régulièrement des modifications pour compenser les effets des changements survenus dans le cadre et les unités de l'échantillon. Ces ajustements peuvent comprendre la création de strates séparées à forte croissance et l'établissement de règles pour faire face aux changements apportés aux divisions administratives qui touchent les unités d'échantillonnage des strates. Cette stratégie de révision a été utilisée pour l'échantillon-maître australien, mais il semble qu'elle soit rarement utilisée dans les pays en développement. La raison est probablement liée à sa complexité du point de vue de l'échantillonnage : elle exige plus de soin et de qualifications dans sa conception et son exécution.

6. Documentation

87. Une grande partie du travail de documentation est déjà effectuée si l'échantillon-maître a été sélectionné à partir d'un cadre directeur d'échantillonnage bien étayé. Cependant, la documentation est parfois l'un des points faibles des échantillons-maîtres dans les pays en développement. Les informations peuvent être éparpillées et parfois rares, de sorte qu'il est difficile de suivre la sélection de l'échantillon et de calculer les probabilités de l'échantillonnage. À chaque stade, pour toutes les unités de l'échantillon-maître, il faut que les procédures de sélection et les probabilités de sélection soient pleinement documentées. Il doit y avoir des pièces indiquant les unités de l'échantillon-maître qui ont été utilisées pour les échantillons de certaines enquêtes. Il convient d'utiliser un système normalisé de numéro d'identification des unités d'échantillonnage.

88. La documentation de l'échantillon-maître devrait donner une évaluation de la performance de l'échantillon-maître mesurée au nombre d'erreurs d'échantillonnage et aux effets liés à la conception pour les estimations importantes. Ces évaluations sont utiles pour la planification de la taille des échantillons et pour l'allocation d'échantillons aux nouvelles enquêtes sur la base de l'échantillon-maître. Les méthodes de calcul des variances et des effets liés à la conception sont à présent décrites dans bon nombre de logiciels d'analyse statistique (voir détails au chapitre XXI).

89. La documentation devrait également comprendre des documents auxiliaires pour l'échantillon-maître. Si des cadres secondaires d'échantillonnage (CSE) ont été préparés pour les unités ultimes de l'échantillon-maître, ces cadres devraient figurer dans la documentation. Les CSE sont composés d'unités géographiques, telles que blocs ou îlots, ou d'unités énumérées, telles qu'unités d'habitation à l'intérieur des unités ultimes de l'échantillon-maître.

7. Utilisation d'un échantillon-maître pour des enquêtes sur des entités économiques

90. L'échantillon-maître a pour principal but de fournir des échantillons pour les enquêtes sur les ménages d'un programme continu d'enquêtes (et pour toute enquête spéciale conforme à la conception de l'échantillon-maître). L'échantillon est donc principalement conçu pour servir un ensemble élémentaire d'enquêtes sur les ménages. En général, il se prête mal à l'échantillonnage d'autres types d'unités. Dans certains cas, cependant, on peut l'utiliser pour des enquêtes portant sur l'étude des caractéristiques de certaines unités économiques, telles qu'entreprises familiales, travailleurs indépendants et petites exploitations agricoles.

91. Dans la plupart des pays développés, un fort pourcentage des entités économiques des secteurs des services, du commerce et de l'agriculture sont étroitement liées à des familles particulières. Ces entités sont généralement nombreuses et de petite taille, et elles sont très répandues parmi la population. Il y a souvent un rapport de un à un entre ces entités et les ménages, et ces derniers plutôt que les entités elles-mêmes peuvent servir d'unités ultimes d'échantillonnage. Un échantillon-maître de ménages peut être utilisé pour les enquêtes sur ces types d'entités. Cela suppose souvent que l'on s'écarte des systèmes d'autopondération de l'échantillon. Verma (2001) examine les moyens d'améliorer l'efficacité de conception des échantillons pour les enquêtes sur les unités économiques.

92. Cela dit, il existe généralement un certain nombre de grands établissements qui ne sont pas liés à une famille. Ces établissements sont généralement peu nombreux, mais ils comptent pour une forte proportion de nombreuses estimations des totaux (production, nombre d'employés, etc.). Souvent, ils sont en outre inégalement répartis par rapport à l'ensemble de la population. Comme l'échantillon-maître des zones ne donne pas un échantillonnage efficace de ces grandes unités, il faut à celles-ci un cadre d'échantillonnage distinct. Dans beaucoup de cas, ce cadre peut être établi à partir des listes des administrations (par exemple, de la direction des impôts ou des services des licences). À partir de ces listes, on choisirait toutes les très grandes unités et un échantillon des autres pour l'enquête, en même temps qu'on prendrait un échantillon d'établissements tirés des UPE de l'échantillon-maître.

93. On se trouve en présence d'un cas particulier d'enquête sur les établissements lorsqu'une enquête sur les ménages est liée à une « enquête communautaire ». Par exemple, dans une enquête sur la santé, l'enquête sur les personnes/ménages pourrait être accompagnée d'une enquête sur les établissements de soins portant sur des zones étendues situées autour de chacune des zones de l'échantillon original (par exemple, les zones d'énumération). Les données recueillies dans le cadre de l'enquête supplémentaire ont deux buts : *a*) elles peuvent être liées aux données sur les ménages et utilisées pour les analyses de la qualité et de l'accessibilité des établissements de soins; et *b*) elles peuvent servir à produire des estimations nationales du nombre et des types d'établissements de soins. Dans le premier cas, les ménages/individus restent l'unité d'analyse : on n'a pas besoin de nouvel échantillonnage. Le second cas est plus compliqué. Si, en même temps que la zone initiale, on prend la zone élargie autour d'elle comme unité (district, commune, zone de supervision du recensement, etc.), composée d'un certain nombre de zones, la situation est simple. L'échantillon produit équivaldrait à un échantillon de zones plus grandes et présentant une probabilité de sélection égale à la somme des probabilités de sélection des zones plus petites composant la zone élargie. En revanche, si la zone élargie est construite selon la règle « dans un rayon de x kilomètres de la zone initiale », la détermination des probabilités de sélection est plus complexe.

E. Conclusion

94. La conception et l'exécution d'enquêtes sur les ménages est une tâche importante pour tous les offices nationaux de statistiques. Nombre d'ONS des pays en développement effectuent plusieurs enquêtes de ce type chaque année. Le besoin de planifier et de coordonner le déroulement de ces enquêtes a stimulé les efforts en faveur de leur intégration dans des programmes intégrés d'enquêtes sur les ménages. L'idée d'un tel programme est aujourd'hui appliquée dans de nombreux offices nationaux de statistiques.

95. Une part importante du travail qu'implique un tel programme est consacrée à la conception d'échantillons pour différentes enquêtes. Ce chapitre a traité les questions clés concernant la conception et l'élaboration de cadres directeurs d'échantillonnage et d'échantillons-maîtres. Il a fait état des avantages qu'offre un cadre directeur d'échantillonnage bien tenu et souligné que tout ONS ayant un programme d'enquêtes sur les ménages devrait avoir un cadre directeur d'échantillonnage bien tenu pour répondre aux besoins des enquêtes de ce programme ainsi qu'aux besoins d'enquêtes spéciales qui peuvent se présenter dans le même temps. Par ailleurs, nombre d'ONS peuvent aller plus loin et concevoir et utiliser un échantillon-maître pour la totalité ou la plupart des enquêtes du programme et, éventuellement, pour les enquêtes spéciales dont le besoin peut apparaître de façon imprévue.

96. Ce chapitre a présenté un aperçu des mesures importantes à prendre lors de l'élaboration de cadres directeurs d'échantillonnage et d'échantillons-maîtres et a donné des illustrations de ces cadres et échantillons dans quelques pays en développement. Son format ne permet pas de traiter de façon détaillée toutes les questions importantes liées à cette élaboration. Les lecteurs qui désireraient en savoir plus sur ces questions sont invités à consulter le manuel pertinent des Nations Unies (voir Nations Unies, 1986).

RÉFÉRENCES

- Ajayi, O. O. (2000). Survey methodology for the sample census of agriculture in Nigeria with some comparisons of experiences in other countries. Étude présentée au Séminaire international sur les résultats du recensement agricole de la Chine tenu à Beijing, 19-22 septembre 2000.
- Glewwe, P., et I.Yansaneh (2001). *Recommendations for Multi-Purpose Household Surveys from 2002 to 2010*. Rapport de mission à l'Office général des statistiques, Viet Nam.
- Kish, L. (1965). *Survey Sampling*. New York, John Wiley and Sons.
- _____ (1988). Multi-purpose sample design. *Survey Methodology*, vol. 14, pp. 19-32.
- Nations Unies (1986). *National Household Survey Capability Programme: Sampling Frames and Sample Designs for Integrated Household Survey Programmes (Preliminary Version)*. DP/UN/INT-84-014/5E, New York.
- Pettersson, H. (1994). *Master Sample Design: Report from a Mission to the National Central Statistics Office, Namibia, May 1994*. International Consulting Office, Statistics Sweden.
- _____ (2001a) *Sample Design for Household and Business Surveys: Report from a Mission to the Bureau of Statistics, Lesotho, 21 May-2 June 2001*. International Consulting Office, Statistics Sweden.
- _____ (2001b). *Recommendations Regarding the Design of a Master Sample for the Household Surveys of GSO: Report of Mission to the General Statistics Office, Viet Nam*. International Consulting Office, Statistics Sweden.

- Rosen, B. (1997). *Creation of the 1997 Lao Master Sample: Report from a Mission to the National Statistics Centre, Lao PDR*. International Consulting Office, Statistics Sweden.
- Torene, R., et L. G. Torene (1987). The practical side of using master samples: the Bangladesh experience. *Bulletin of the International Statistical Institute: Proceedings of the 46th Session, Tokyo, 1987*, vol. LII-2, pp. 493-511.
- Verma, V. (2001). Sample design for national surveys: surveying small-scale economic units. *Statistics in Transition*, vol. 5, n° 3 (décembre 2001), pp. 367-382.

Chapitre VI

Estimation des composantes des effets liés à la conception à utiliser lors de la conception d'échantillons

GRAHAM KALTON
Westat
Rockville, Maryland
États-Unis d'Amérique

J. MICHAEL BRICK
Westat
Rockville, Maryland
États-Unis d'Amérique

THANH LÊ
Westat
Rockville, Maryland
États-Unis d'Amérique

RÉSUMÉ

L'effet lié à la conception – le ratio de la variance d'une statistique obtenue à l'aide d'un échantillon complexe à la variance d'une statistique obtenue à l'aide d'un échantillon aléatoire simple ou d'un échantillon de même taille sans restriction – est un outil précieux pour la conception d'un échantillon. Cependant, un effet lié à la conception constaté dans une enquête ne doit pas être automatiquement adopté pour être utilisé pour la conception d'une autre enquête. L'effet lié à la conception représente l'effet combiné d'un certain nombre de composantes telles que la stratification, la mise en grappes, une sélection inégale de probabilités et les ajustements de pondération pour non-réponse et non-couverture. Plutôt que de se contenter d'importer un effet global lié à la conception d'une enquête précédente, il conviendrait d'examiner soigneusement les diverses composantes en jeu. Le présent chapitre est consacré à l'examen des effets liés à la conception dus aux diverses composantes, puis à décrire les modèles qui peuvent être utilisés pour combiner ces effets en un effet lié à la conception globale. À partir de ces composantes, le concepteur de l'échantillon peut établir des estimations des effets globaux liés à la conception pour divers modèles possibles d'échantillon, puis utiliser ces estimations pour choisir une conception d'échantillon efficace pour l'enquête prévue.

Termes clés : stratification, mise en grappes, pondération, coefficient de corrélation à l'intérieur d'une classe.

A. Introduction

1. Comme le montrent d'autres chapitres de la présente publication, les enquêtes nationales sur les ménages dans les pays en développement et les pays en transition emploient des méthodes complexes d'échantillonnage, y compris l'échantillonnage à phases multiples, la stratification et, souvent, des probabilités de sélection inégales. L'une des conséquences de cette situation est que l'on ne peut calculer les erreurs d'échantillonnage des estimations d'une enquête en utilisant les formules que l'on trouve ordinairement dans les manuels de statistiques. Ces formules sont fondées sur l'hypothèse selon laquelle les variables observées sont des variables aléatoires réparties

de façon indépendante et identique [*independently and identically distributed (iid)*]. Cette hypothèse ne tient pas pour les observations découlant d'une conception complexe de l'échantillon, de sorte qu'il faut recourir à une approche différente de l'estimation des erreurs d'échantillonnage des estimations de l'enquête.

2. On peut calculer les variances des estimations d'une enquête établies à partir d'échantillons de conception complexe en employant un certain type de méthode de réplication, telle que la méthode appelée « jackknife repeated replication » ou « balanced repeated replication », ou la méthode appelée « Taylor series linearization method » [voir, par exemple, Wolter (1985); Rust (1985); Verma (1996); Lehtonen et Pahkinen (1994); Rust et Rao (1996)]. Il existe un certain nombre de programmes informatiques spécialisés pour effectuer ces calculs [voir l'analyse d'un grand nombre de ces programmes donnée par Lepkowski et Bowles (1996), également accessible à <http://www.fas.harvard.edu/~stats/survey-soft/iass.html>; et l'aperçu donné du logiciel d'analyse des enquêtes préparé par la Section des méthodes de recherche sur les enquêtes de l'Association américaine de statistiques, accessible à <http://www.fas.harvard.edu/~stats/survey-soft/survey-soft.html>]. Lorsque les variances sont calculées d'une manière qui tient compte de la complexité de conception de l'échantillon, leurs estimations diffèrent de celles qui seraient obtenues par application des formules types fondées sur l'utilisation de variables *iid*. Dans bien des cas, les variances liées à une conception complexe sont plus grandes, souvent dans des proportions sensibles, que celles obtenues à l'aide des formules types.

3. Les formules de calcul des variances que l'on trouve dans les manuels de statistiques sont applicables à une forme de conception d'échantillon, à savoir l'échantillonnage sans restrictions (également connu sous le nom d'échantillonnage aléatoire simple avec remplacement). Selon cette conception, les unités de la population de l'enquête sont sélectionnées de façon indépendante et avec une probabilité égale. Ces unités sont sélectionnées avec remplacement, ce qui implique qu'une unité peut figurer plusieurs fois dans l'échantillon. Supposons qu'un échantillon sans restrictions de taille n produit des valeurs Y_1, Y_2, \dots, Y_n pour la variable y . La variance de la moyenne de l'échantillon $\bar{y} = \sum y_i / n$ est :

$$V_u(\bar{y}) = \sigma^2 / n \quad (1)$$

où $\sigma^2 = \sum^N (Y_i - \bar{Y})^2 / N$ est la variance élémentaire des N valeurs- y de la population (Y_1, Y_2, \dots, Y_N) et $\bar{Y} = \sum Y_i / N$. On peut estimer cette variance à partir de l'échantillon selon la formule :

$$v_u(\bar{y}) = s^2 / n \quad (2)$$

où $s^2 = \sum^n (y_i - \bar{y})^2 / (n - 1)$. On trouve cette même formule dans les manuels de statistiques courants.

4. En règle générale, les échantillons sont sélectionnés de préférence sans remplacement, car les estimations de l'enquête sont alors plus précises (autrement dit, elles ont des variances plus faibles) lorsque des unités ne peuvent être incluses qu'une seule fois dans l'échantillon. Avec l'échantillonnage aléatoire simple sans remplacement, généralement dénommé échantillonnage aléatoire simple (*simple random sampling* ou SRS), les unités sont sélectionnées avec la même probabilité, et tous les ensembles possibles d'unités distinctes n de la population de N unités ont la même probabilité de figurer dans l'échantillon. Avec un SRS de taille n , la variance et l'estimation de la variance de la moyenne de l'échantillon $\bar{y} = \sum y_i / n$ sont données par :

$$V_0(\bar{y}) = (1 - f)S^2 / n \quad (3)$$

et

$$v_0(\bar{y}) = (1 - f)s^2 / n \tag{4}$$

où $f = n/N$ est la fraction d'échantillonnage, $S^2 = \sum^N (Y_i - \bar{Y})^2 / (N - 1)$ et $s^2 = \sum^n (y_i - \bar{y})^2 / (n - 1)$. Lorsque N est élevé, comme c'est généralement le cas dans la recherche par enquête, σ^2 et S^2 sont sensiblement égaux. Autrement dit, la principale différence entre la variance pour la moyenne de l'échantillonnage sans restrictions dans l'équation (1) et la variance pour le SRS dans (3) est le facteur $(1 - f)$, connu sous le nom de correction en population finie (*fpc*). Dans la plupart des situations pratiques, la fraction d'échantillonnage n/N est petite et peut être traitée comme égale à 0. Lorsque tel est le cas, le terme *fpc* de (3) et (4) est approximativement égal à l'unité, et la distinction entre l'échantillonnage avec et sans remplacement peut être ignorée.

5. Les formules de calcul de la variance données ci-dessus ne peuvent être appliquées aux échantillons de conception complexe, mais elles peuvent servir de critères de comparaison pour les variances des estimations de tels échantillons. Kish (1965) a inventé l'expression « *design effect* » (effet lié à la conception) pour indiquer le ratio de la variance de toute estimation, disons z , obtenu à partir d'une conception complexe à la variance de z applicable à un SRS ou à un échantillon sans restrictions de la même taille¹. Il convient de noter que l'effet lié à la conception se rapporte à une estimation donnée de l'enquête, et diffère pour différentes estimations d'une enquête donnée. Il convient également de noter que z peut être n'importe quelle estimation recherchée, par exemple une moyenne, une proportion, un total, ou un coefficient de régression.

6. L'effet lié à la conception à la fois de la forme de l'échantillon de conception complexe employé et de l'estimation de l'enquête examinée. Pour incorporer ces deux caractéristiques, nous employons la notation $D^2(z)$ pour l'effet lié à la conception de l'estimation z , où

$$D^2(z) = \frac{\text{Variance avec une conception complexe}}{\text{Variance de } z \text{ avec une conception d'échantillon sans restrictions de même taille}} = \frac{V_c(z)}{V_u(z)} \tag{5}$$

Le carré est employé pour permettre l'utilisation de $D(z)$ comme la racine carrée de l'effet lié à la conception. La simple mention de $D(z)$ est utile car ce terme représente le multiplicateur à appliquer à l'erreur type de z dans le cas d'une conception d'échantillon sans restrictions pour indiquer son erreur type pour une telle conception comme, par exemple, le calcul de l'intervalle de confiance.

7. La « taille effective de l'échantillon », indiquée n_{eff} , est un concept utile en rapport direct avec l'effet lié à la conception. La taille effective de l'échantillon est celle d'un échantillon sans restrictions qui donnerait le même niveau de précision à l'estimation de l'enquête qu'un échantillon de conception complexe. Elles est donc indiquée par :

$$n_{eff} = n / D^2(z) \tag{6}$$

8. La définition de $D^2(z)$ donnée ci-dessus est valable pour un travail théorique où les vraies variances $V_c(z)$ et $V_u(z)$ sont connues. Dans les applications pratiques, ces variances sont estimées à partir de l'échantillon, et $D^2(z)$ est alors estimé par $d^2(z)$. Ainsi,

$$d^2(z) = \frac{v_c(z)}{v_u(z)} \tag{7}$$

où $v_c(z)$ est estimée selon une procédure qui convient à la conception complexe et $v_u(z)$ est estimée selon une formule applicable à l'échantillonnage sans restrictions,

¹ Plus précisément, Kish (1982) a défini *Deff* comme ce ratio avec un dénominateur de la variance du SRS, et *Deft*² comme le ratio avec un dénominateur de la variance de l'échantillon sans restrictions. La différence entre *Deff* et *Deft*² dépend de la question de savoir si le terme *fpc* $(1 - f)$ est ou non inclus. Comme ce terme a un effet négligeable dans la plupart des enquêtes nationales sur les ménages, la distinction entre *Deff* et *Deft*² est rarement significative et sera donc ignorée dans le reste de ce chapitre. Tout au long de cet ouvrage, nous supposons que le terme *fpc* peut être ignoré. Voir également Kish (1995). Skinner a défini un concept différent mais apparenté, l'effet lié à la méspécification ou *meff*, qui, selon lui, se prête mieux à l'analyse des données de l'enquête (voir, par exemple, Skinner, Holt et Smith (1989), chap. 2). Comme ce chapitre traite de la conception de l'échantillon plutôt que de l'analyse, ce concept ne sera pas discuté ici. Cette référence résume un grand nombre de résultats sous une forme très utile. Beaucoup de ces relations sont connues et ont été publiées il y a plusieurs dizaines d'années. Voir, par exemple, Kish (1965) et Kish (1976).

et les paramètres inconnus sont estimés à partir de l'échantillon. Ainsi, par exemple, dans le cas de l'échantillon moyen :

$$v_u(z) = s^2/n \quad (8)$$

et pour les grands échantillons, s^2 peut être estimé par :

$$\frac{\sum w_i (y_i - \bar{y})^2}{\sum w_i}$$

où y_i et w_i sont la valeur- y et la pondération de l'unité échantillonnée i et $\bar{y} = \sum w_i y_i / \sum w_i$ est l'estimation pondérée de la moyenne de la population. Dans le cas d'une proportion d'échantillon p , pour un grand n :

$$v_u(p) = \frac{p(1-p)}{n-1}$$

ou

$$v_u(p) = \frac{p(1-p)}{n}$$

où p est l'estimation pondérée de la proportion de la population.

9. Une autre question doit être traitée en ce qui concerne la définition des effets liés à la conception et leur estimation. De nombreuses enquêtes emploient des modèles d'échantillons qui présentent des probabilités de sélection inégales, de sorte que certains sous-groupes risquent d'être représentés de façon disproportionnée dans l'échantillon. Par exemple, il peut arriver que, dans une enquête nationale sur les ménages, 50 % de l'échantillon de 2 000 ménages soient prélevés dans les zones urbaines et 50 % dans les zones rurales alors que seuls 30 % des ménages vivent en milieu urbain. Considérons, par exemple, l'effet lié à la conception pour une moyenne estimative des ménages urbains. Le dénominateur de (8) est s^2/n . La question est de savoir comment calculer n . Une solution peut être d'utiliser dans ce cas la taille effective de l'échantillon urbain, 1 000. On pourrait aussi utiliser la taille attendue de l'échantillon dans les zones urbaines pour un SRS de $n = 2 000$, qui est ici de $0,3 \times 2 000 = 600$. La première de ces deux approches, qui part de l'échantillon effectif de 1 000, est celle qui est le plus couramment utilisée, et c'est celle qui sera utilisée dans ce chapitre. Cependant, certains programmes d'estimation de la variance permettent de calculer les effets liés à la conception selon la seconde approche. Comme ces deux approches peuvent produire des valeurs sensiblement différentes, il importe de les distinguer et de choisir l'option appropriée.

10. La notion d'effet lié à la conception s'est révélée utile pour la conception d'échantillons simples. Les échantillons complexes nécessitent une combinaison de plusieurs composantes de conception, telles que la stratification, l'échantillonnage à phases multiples et la sélection assortie de probabilités inégales. L'analyse des effets liés à la conception pour chacune de ces composantes nous éclaire sur leurs effets sur la précision des estimations de l'enquête et facilite donc l'élaboration de modèles d'échantillon efficaces. Nous examinerons les effets liés à la conception pour ces diverses composantes à la section B. Lors de la conception d'un échantillon complexe, il est utile de construire des modèles qui permettent de prédire l'ensemble des effets sur la conception découlant d'une combinaison de composantes. Nous examinerons brièvement ces modèles à la section C. Nous présenterons à titre d'illustration un exemple hypothétique de l'utilisation des effets liés à la conception pour la conception d'un

échantillon à la section D, et concluons par quelques observations générales à la section E.

B. Composantes des effets liés à la conception

11. La présente section traite des effets liés à la conception résultant des composantes suivantes d'un échantillon de conception complexe : stratification proportionnée et disproportionnée; mise en grappes; probabilités de sélection inégales; ajustements de pondération pour tenir compte des cas de non-réponse; et ajustements de pondération de la population pour tenir compte des cas de non-couverture et pour améliorer la précision. Ces diverses composantes sont examinées séparément dans cette section; leurs effets combinés sont traités à la section C. La principale statistique considérée est une estimation d'une moyenne concernant la population (par exemple, le revenu moyen). Comme une proportion de la population (par exemple, la proportion de la population vivant en situation de pauvreté) est en fait un cas particulier d'une moyenne arithmétique, ce traitement couvre aussi une proportion. Les proportions sont probablement les statistiques les plus largement utilisées dans les rapports d'enquête, et seront donc traitées séparément en temps voulu. De nombreux résultats d'enquête se rapportent à des sous-groupes de la population totale, tels que les femmes âgées de 15 à 44 ans, ou aux personnes vivant dans les zones rurales. Nous examinerons donc les effets de la pondération et de la mise en grappes sur les estimations relatives à des sous-groupes.

1. Stratification

12. Nous commencerons par examiner l'effet lié à la conception pour la moyenne d'un échantillon stratifié à phase unique avec échantillonnage aléatoire simple à l'intérieur des strates. La moyenne de l'échantillon stratifié est donnée par :

$$\bar{y}_{st} = \sum_h \frac{N_h}{N} \sum_i \frac{y_{hi}}{n_h} = \sum_h W_h \bar{y}_h$$

où n_h est la taille de l'échantillon choisi parmi les N_h unités de la strate h , $N = \sum N_h$ est la taille de la population, $W_h = N_h/N$ est la proportion de la population de la strate h , y_{hi} est la valeur de l'unité sélectionnée i de la strate h , et $\bar{y}_h = \sum_i y_{hi} / n_h$ est la moyenne de l'échantillon de la strate h . Dans la pratique, \bar{y}_{st} se calcule comme une estimation pondérée, où l'on affecte à chaque unité sélectionnée une pondération de base qui est l'inverse de sa probabilité de sélection (on ne tient pas compte pour le moment des ajustements de pondération de l'échantillon et de la population). Ici, chaque unité de la strate h a une probabilité de sélection de n_h/N_h et par conséquent une pondération de base de $w_{hi} = w_h = N_h/n_h$. Ainsi, \bar{y}_{st} peut s'exprimer comme :

$$y_{st} = \frac{\sum_h \sum_i w_{hi} y_{hi}}{\sum_h \sum_i w_{hi}} = \frac{\sum_h \sum_i w_h y_{hi}}{\sum_h n_h w_h} \quad (9)$$

Si l'on suppose que l'on peut ne pas tenir compte de la correction en population finie, la variance de la moyenne stratifiée est donnée par :

$$V(\bar{y}_{st}) = \sum_h \frac{W_h^2 S_h^2}{n_h} \quad (10)$$

où $S_h^2 = \sum_i (Y_{hi} - \bar{Y}_h)^2 / (N_h - 1)$ est la variance de l'unité de population de la strate h .

13. L'ampleur de $V(\bar{y}_{st})$ dépend de la façon dont l'échantillon est distribué à travers la strate. Dans le cas courant où l'on utilise une allocation proportionnée, de façon que la taille de l'échantillon à l'intérieur d'une strate soit proportionnelle à la taille de la population de cette strate, les pondérations de toutes les unités sélectionnées sont les mêmes. La moyenne stratifiée est réduite à la simple moyenne non pondérée $\bar{y}_{prop} = \sum \sum y_{hi} / n$, où $n = \sum n_h$ est la taille globale de l'échantillon, et sa variance est réduite à :

$$V(\bar{y}_{prop}) = \frac{\sum W_h S_h^2}{n} = \frac{S_w^2}{n} \quad (11)$$

où S_w^2 indique la variance moyenne de l'unité à l'intérieur de la strate. L'effet lié à la conception pour \bar{y}_{prop} pour un échantillon stratifié proportionné s'obtient donc en utilisant la variance de la moyenne d'un échantillon aléatoire simple de l'équation (3), sans tenir compte du terme *fpc*, et la définition de l'effet lié à la conception de l'équation (5) :

$$D^2(\bar{y}_{prop}) = \frac{S_w^2}{S^2} \quad (12)$$

Comme la variance moyenne de l'unité à l'intérieur d'une strate n'est pas supérieure à la variante globale de l'unité (à condition que les valeurs de N_h soient élevées), l'effet lié à la conception pour la moyenne d'un échantillon proportionné n'est pas supérieur à 1. Autrement dit, la stratification proportionnée ne peut entraîner une perte de précision et aboutit généralement, au contraire, à un gain de précision. Celui-ci se produit lorsque les moyennes des strates Y_h diffèrent : plus la variation entre les strates est grande, plus le gain est élevé.

14. Dans beaucoup d'enquêtes, on a besoin d'un échantillon stratifié disproportionné pour que l'enquête puisse donner des estimations pour certains domaines particuliers. Par exemple, l'enquête peut avoir pour objectif de produire des estimations fiables pour chaque région d'un pays où la population varie d'une région à l'autre. Pour ce faire, il peut être nécessaire d'allouer aux petites régions des échantillons dont la taille soit sensiblement plus grande que celle de l'échantillon qui leur serait alloué par un échantillonnage stratifié proportionnel. Les coûts de collecte de données qui diffèrent sensiblement selon les strates peuvent offrir une autre raison de s'écarter de l'allocation proportionnelle. Dans ce cas, la formule optimale serait d'allouer des échantillons plus grands qu'avec une allocation proportionnelle aux strates où les coûts de collecte des données sont plus faibles.

15. Le gain de précision découlant d'une stratification proportionnée ne s'applique pas nécessairement au cas d'une allocation disproportionnée de l'échantillon. Pour simplifier le débat sur ce point, nous supposons que les variances de la population à l'intérieur d'une strate sont constantes, autrement dit que $S_h^2 = S_c^2$ pour toutes les strates. Cette supposition est souvent raisonnable pour les enquêtes nationales sur les ménages où la stratification disproportionnée est utilisée pour les raisons indiquées ci-dessus. Selon cette supposition, l'équation (10) est simplifiée comme suit :

$$V(\bar{y}_{st}) = S_c^2 \sum_h \frac{W_h^2}{n_h} = \frac{S_c^2}{N} \sum_h W_h w_h \quad (13)$$

Dans ce cas, l'effet lié à la conception est :

$$D^2(\bar{y}_{st}) = \frac{S_c^2}{S^2} = \frac{n}{N} \sum_h W_h w_h \quad (14)$$

16. Outre le fait que nous supposons des variances constantes à l'intérieur des strates pour parvenir à l'équation (14), il est souvent raisonnable de supposer que les moyennes des strates sont plus ou moins égales, c'est-à-dire que $\bar{Y}_h = \bar{Y}$ pour toutes les strates. Avec cette nouvelle supposition, $S_c^2 = S^2$. et l'effet lié à la conception se ramène à :

$$D^2(\bar{y}_{st}) = \frac{n}{N} \sum_h W_h w_h = n \sum_h \frac{W_h^2}{n_h} \quad (15)$$

Kish (1992)² présente l'effet lié à la conception dû à une allocation disproportionnée comme :

$$D^2(\bar{y}_{st}) = (\sum_h W_h w_h) (\sum_h W_h / w_h) \quad (16)$$

Cette formule est très utile pour la conception de l'échantillon. Toutefois, on ne saurait l'appliquer aveuglément sans tenir compte du caractère raisonnable des hypothèses sous-jacentes (voir ci-dessous).

17. Comme simple exemple de l'application de l'équation (16), considérons un pays ayant deux régions dont la première contient 80 % de la population totale et la seconde 20 % (autrement dit, $W_1 = 4W_2$). Supposons qu'une enquête est effectuée avec des échantillons de même taille pour les deux régions ($n_1 = n_2 = 1\,000$). On peut utiliser l'une quelconque des expressions ci-dessus pour calculer l'effet lié à la conception à partir d'une allocation disproportionnée de la moyenne estimative nationale (en supposant que la variance moyenne et la variance unitaire soient les mêmes dans les deux régions). Par exemple, si l'on utilise l'équation (16) et si l'on tient compte du fait que $w_1 = 4w_2$, l'effet lié à la conception est :

$$D_w^2(\bar{y}_{st}) = (4W_2 \cdot 4w_2 + W_2 \cdot w_2) \left(\frac{4W_2}{4w_2} + \frac{W_2}{w_2} \right) = 1,36$$

car $W_2 = 0,2$. L'allocation disproportionnée utilisée pour obtenir sensiblement la même précision pour les estimations de chaque région se traduit par une moyenne estimative pour l'ensemble du pays avec un échantillon effectif de taille $n_{eff} = 2\,000/1,36 = 1\,471$.

18. Le tableau VI.1 illustre l'effet lié à la conception dû à une allocation disproportionnée pour certains taux de suréchantillonnage couramment utilisés lorsqu'il n'y a que deux strates. Les chiffres figurant en tête de chaque colonne sont les ratios des pondérations dans les deux strates, qui sont équivalents à l'inverse des ratios des taux d'échantillonnage dans les deux strates. Les sous-éléments sont les proportions de la population de la première strate. Comme l'effet lié à la conception est symétrique autour de 0,50, on peut obtenir les valeurs pour $W_1 > 0,5$ en utilisant la rangée correspondant à $(1 - W_1)$. Pour illustrer l'utilisation de ce tableau, considérons l'exemple donné ci-dessus. La valeur dans la rangée où $W_1 = 0,20$ et la colonne où le ratio de suréchantillonnage est 4 donne $D^2(\bar{y}_{st}) = 1,36$. Le tableau montre que les effets liés à la conception augmentent à mesure que le ratio des taux d'échantillonnage augmente et que la proportion de la population de la strate approche les 50 %. Lorsque les taux d'échantillonnage sont très différents entre les strates, l'effet lié à la conception pour la moyenne globale peut être très prononcé et, par conséquent, l'échantillon est de petite

² Cette référence résume un grand nombre de résultats sous une forme très utile. Un grand nombre de ces relations sont connues et ont été publiées il y a plusieurs dizaines d'années. Voir, par exemple, Kish (1965) et Kish (1976).

taille. Dans ce cas, l'allocation disproportionnée engendre un échantillon très inefficace pour estimer la statistique applicable à l'ensemble de la population.

19. De nombreuses enquêtes nationales sont conçues pour produire des estimations nationales mais aussi des estimations pour les diverses régions du pays. Généralement, les régions sont de tailles très variables. En pareil cas, on se heurte à un conflit quand on cherche à déterminer une allocation appropriée de l'échantillon à toutes les régions, comme le montrent les résultats ci-dessus. Selon les hypothèses de variances moyennes et de variances unitaires égales à l'intérieur des régions, l'allocation optimale pour les estimations nationales est une allocation proportionnée, tandis que pour les estimations régionales, c'est une taille égale de l'échantillon pour chaque région. L'utilisation de l'allocation optimale dans un cas se traduit par un mauvais échantillon dans l'autre. Néanmoins, une allocation de compromis peut donner des résultats raisonnablement acceptables dans les deux cas (voir section D).

Tableau VI.1

Effets liés à la conception d'un échantillonnage disproportionné dans l'hypothèse de deux strates

W_1	Ratio de w_1 à w_2								
	1	2	3	4	5	8	10	20	
0,05	1,00	1,02	1,06	1,11	1,15	1,29	1,38	1,86	
0,10	1,00	1,05	1,12	1,20	1,29	1,55	1,73	2,62	
0,15	1,00	1,06	1,17	1,29	1,41	1,78	2,03	3,30	
0,20	1,00	1,08	1,21	1,36	1,51	1,98	2,30	3,89	
0,25	1,00	1,09	1,25	1,42	1,60	2,15	2,52	4,38	
0,35	1,00	1,11	1,30	1,51	1,73	2,39	2,84	5,11	
0,50	1,00	1,13	1,33	1,56	1,80	2,53	3,03	5,51	

20. L'équation (16) est amplement utilisée dans la conception de l'échantillon pour évaluer l'effet de l'utilisation d'une allocation disproportionnée sur les estimations nationales. Toutefois, en l'utilisant, les utilisateurs devraient tenir compte des hypothèses de moyennes et de variances égales à l'intérieur des strates sur lesquelles elle se fonde. Considérons la première situation où les moyennes sont différentes mais pas les variances. Dans ce cas, l'effet lié à la conception résultant d'une stratification disproportionnée est donné par l'équation (14), avec le facteur additionnel $S_c^2 = S^2$. Ce facteur est inférieur à l'unité, de sorte que l'effet lié à la conception n'est pas aussi grand que celui donné par l'équation (16). Toutefois, cet effet représente l'effet global de la stratification et de l'allocation disproportionnée. Pour ne mesurer que l'effet de l'allocation disproportionnée, la comparaison à faire est entre l'échantillon stratifié disproportionné et un échantillon stratifié proportionné de même taille. Le ratio de la variance de \bar{y}_{st} pour l'échantillon disproportionné à celle de \bar{y}_{prop} est, d'après les équations (11) et (13) avec $S_w^2 = S_c^2$:

$$R = V(\bar{y}_{st}) / V(\bar{y}_{prop}) = (\sum_h W_h w_h) (\sum_h W_h / w_h)$$

Ainsi, dans ce cas, la formule de l'équation (16) peut être interprétée comme l'effet de la seule allocation disproportionnée.

21. L'hypothèse de variances unitaires égales à l'intérieur des strates est plus délicate. Il ressort des résultats ci-dessus qu'une allocation disproportionnée entraîne

une perte de précision des estimations globales lorsque les variances unitaires à l'intérieur d'une strate sont égales, mais tel n'est pas nécessairement le cas lorsque ces variances unitaires à l'intérieur d'une strate sont inégales. En fait, dans ce cas, les fractions d'échantillonnage optimales à utiliser sont proportionnelles à l'écart type à l'intérieur des strates [voir, par exemple, Cochran (1977)]. Ce type d'allocation disproportionnée est largement utilisé dans les enquêtes sur les entreprises. Il peut se traduire par des gains de précision substantiels par rapport à une allocation proportionnée lorsque les écarts types à l'intérieur des strates sont sensiblement différents.

22. Dans les enquêtes sur les ménages, l'hypothèse de variances égales ou plus ou moins égales à l'intérieur des strates est souvent raisonnable. Le type d'estimations pour lesquelles les variances à l'intérieur des strates peuvent être inégales est la proportion. La proportion est la moyenne d'une variable qui n'a pour valeurs que 1 et 0, signifiant la présence ou l'absence d'une caractéristique donnée. La variance unitaire pour une telle variable est $\sigma^2 = P(1 - P)$, où P est la proportion de la population qui présente cette caractéristique. Ainsi, la variance unitaire à l'intérieur de la strate h avec une proportion P_h présentant la caractéristique est $S_h^2 = P_h(1 - P_h)$. Si P_h varie à travers les strates, S_h^2 varie également. Toutefois, la variation de S_h^2 est faible pour les proportions comprises entre 0,2 et 0,8, allant d'un niveau plus élevé de 0,25 pour $P_h = 0,5$ à une valeur plus faible de 0,16 pour $P_h = 0,2$ ou 0,8.

23. Pour illustrer l'effet de la variabilité des proportions et, par conséquent, les variances à l'intérieur d'une strate, nous revenons à notre exemple de deux strates avec $W_1 = 0,8$, $W_2 = 0,2$, et $n_1 = n_2$, et considérons deux ensembles de valeurs différentes pour P_1 et P_2 . Pour le cas 1, supposons que $P_1 = 0,5$ et $P_2 = 0,8$. Dès lors, l'effet global lié à la conception, calculé à l'aide des équations (10) et (1), est $D^2(\bar{y}_{st}) = 1,35$ et le ratio des variances pour les allocations disproportionnée et proportionnée est $R = 1,43$. Pour le cas 2, supposons que $P_1 = 0,8$ et $P_2 = 0,5$. Dès lors, $D^2(\bar{y}_{st}) = 1,16$ et $R = 1,26$. Les valeurs obtenues pour $D^2(\bar{y}_{st})$ et R dans ces deux cas peuvent se comparer à l'effet lié à la conception de 1,36 qui a été obtenu dans l'hypothèse de variances égales à l'intérieur des strates. Dans les deux cas, les effets globaux liés à la conception sont inférieurs à 1,36 en raison du gain de précision découlant de la stratification. Dans le cas 1, la valeur de R est supérieure à 1,36, car la strate 1, où l'échantillon est prélevé à un taux moins élevé, a la variance la plus grande à l'intérieur de la strate. Dans le cas 2, c'est l'inverse : la strate 2, qui est suréchantillonnée, a la plus faible variance. Ce suréchantillonnage est donc la voie à suivre pour accroître la précision. En fait, dans ce cas, l'allocation optimale serait de prélever l'échantillon de la strate 2 à un taux de 1,25 fois plus élevé que celui de la strate 1. Même si les proportions entre les strates sont très différentes dans ces exemples et, par conséquent, les variances à l'intérieur des strates diffèrent aussi sensiblement, les valeurs de R obtenues — à 1,26 et 1,43 — sont raisonnablement proches de 1,36. Ces calculs illustrent le fait que la mesure approximative de l'effet lié à la conception découlant des pondérations produites par l'équation (16) convient à la plupart des cas de planification, même lorsque les variances à l'intérieur des strates sont sensiblement différentes.

24. Enfin, considérons un exemple plus extrême avec $P_1 = 0,05$ et $P_2 = 0,5$, mais toujours avec $W_1 = 0,8$, $W_2 = 0,2$ et $n_1 = n_2$. Dans ce cas, $D^2(\bar{y}_{st}) = 0,67$ et $R = 0,92$. Cet exemple montre qu'une stratification disproportionnée peut engendrer des gains de précision. Toutefois, étant donné les hypothèses sur lesquelles elle est fondée, l'équation (16) ne peut produire une valeur inférieure à 1. Par conséquent, l'équation (16) ne devrait pas être utilisée aveuglément sans tenir compte des hypothèses sous-jacentes.

2. Mise en grappes

25. Nous examinerons à présent une autre composante importante de l'effet global lié à la conception dans la plupart des enquêtes générales sur la population, à savoir l'effet lié à la conception, produit par la mise en grappes d'échantillons à phases multiples. Les échantillons sont mis en grappes pour réduire les coûts de collecte des données car il n'est pas économique d'établir une liste et de sélectionner des ménages dispersés sur l'ensemble du territoire d'un pays ou d'une région. On procède donc à un échantillonnage en deux phases ou plus, où les unités de la première phase ou unités primaires d'échantillonnage (UPE) sont des zones géographiques clairement définies où l'échantillonnage se fait généralement selon des probabilités proportionnelles aux nombres estimatifs des ménages ou des personnes qui les composent. À l'intérieur des UPE sélectionnées, il peut être procédé à une ou plusieurs phases additionnelles d'échantillonnage puis, dans les sous-régions finalement sélectionnées, les unités d'habitation sont énumérées et des ménages sont sélectionnés à partir des listes établies. Pour une enquête sur les ménages, les données sont recueillies pour les ménages sélectionnés. Pour une enquête sur les personnes, on établit la liste des personnes composant les ménages sélectionnés, et l'on sélectionne la totalité ou un échantillon de ces personnes pour l'enquête. Aux fins de la présente analyse, nous supposons qu'il s'agit d'une enquête sur les ménages ne comportant que deux phases d'échantillonnage (les UPE et les ménages). Néanmoins, l'extrapolation aux enquêtes à phases multiples est directe.

26. Dans la réalité, la taille des UPE (c'est-à-dire en ce qui concerne le nombre d'unités qui les composent) est toujours variable et c'est pourquoi les UPE sont sélectionnées selon une probabilité proportionnelle à leur taille estimative [*probability proportional to estimated size* (PPES)]. Les tailles des échantillons sélectionnés à partir de ces UPE varient elles aussi selon les UPE. Toutefois, pour des raisons de simplicité, nous commencerons par supposer que la population est composée de A UPE (par exemple, les districts d'énumération du recensement) dont chacune comprend B ménages. Un simple échantillon aléatoire de a UPE est sélectionné, de même qu'un simple échantillon aléatoire de $b \leq B$ ménages dans chaque UPE sélectionnée (le cas particulier où $b = B$ représente un échantillon en grappes à une seule phase). Nous supposons que le facteur de correction pour population finie de la première phase est négligeable. La conception de l'échantillon pour la sélection des ménages utilise la méthode de sélection à probabilité égale [*equal probability of selection method* (*epsem*)], de sorte que la moyenne de la population peut être estimée par la simple moyenne non pondérée de l'échantillon $\bar{y}_{cl} = \sum_{\alpha}^a \sum_{\beta}^b y_{\alpha\beta} / n$, où $n = ab$ et l'indice cl indique la grappe. La variance de \bar{y}_{cl} est donnée par l'équation :

$$V(\bar{y}_{cl}) = \frac{S^2}{n} [1 + (b-1)\rho] \quad (17)$$

où S^2 est la variance unitaire de la population et ρ est le coefficient de corrélation interne qui mesure l'homogénéité de la variable y des UPE. Dans la pratique, les unités d'une UPE tendent à être plus ou moins semblables les unes aux autres pour presque toutes les variables, quoique le degré de similarité soit généralement faible. Par conséquent, ρ est presque toujours positif et faible.

27. Dans cette situation simple, l'effet lié à la conception est :

$$D^2(\bar{y}_{cl}) = 1 + (b-1)\rho \quad (18)$$

Ce résultat élémentaire montre que l'effet lié à la conception découlant de la mise en grappes de l'échantillon à l'intérieur des UPE dépend de deux facteurs : la taille des

sous-échantillons à l'intérieur des UPE sélectionnées (b) et la corrélation à l'intérieur d'une même classe (ρ). Comme ρ est généralement positif, en règle générale, l'effet lié à la conception découlant de la mise en grappes est supérieur à 1.

28. L'un des traits importants de l'équation (18), et d'autres semblables présentées ci-dessous, est qu'elle est fonction de ρ qui mesure l'homogénéité des UPE pour une variable donnée³. La valeur de ρ est proche de zéro pour beaucoup de variables (par exemple, l'âge et le sexe), et faible mais non négligeable pour d'autres (par exemple, $\rho = 0,03$ à $0,05$), mais elle peut être élevée pour certaines (par exemple, l'accès à une clinique dans le village, l'UPE, quand tous les habitants d'un village ont ou n'ont pas un tel accès). Il est théoriquement possible à ρ d'être négatif, mais cela a peu de chances de se produire dans la réalité (quoique les estimations de ρ soient souvent négatives dans les échantillons). Souvent, ρ est inversement fonction de la taille de l'UPE car les grappes plus grandes ont tendance à être plus diverses, surtout lorsque les UPE sont des zones géographiques. Ces types de relations sont exploitées pour la conception optimale des enquêtes, et les UPE grandes et plus diverses sont utilisées quand il y a la possibilité de choisir. On a besoin d'estimations de ρ pour certaines variables clés de l'enquête pour la conception des échantillons. Ces estimations sont généralement fondées sur les estimations d'enquêtes précédentes pour les mêmes variables et UPE ou pour des variables et UPE similaires, et sur la conviction de portabilité des valeurs de ρ à travers des variables et UPE similaires.

29. Dans la réalité, les UPE ne sont pas de taille égale et elles ne sont pas sélectionnées par échantillonnage aléatoire simple. Dans la plupart des modèles nationaux d'échantillons de ménages, on utilise des échantillons stratifiés d'UPE avec l'échantillonnage PPES. De ce fait, l'équation (18) n'est pas directement applicable. Toutefois, elle reste un modèle utile pour l'effet lié à la conception résultant de la mise en grappes d'une variété d'échantillons *epsem* avec une modification appropriée quant à l'interprétation de ρ .

30. Considérons d'abord un échantillon PPS d'UPE où les mesures exactes de la taille sont connues. Dans ce cas, la combinaison d'un échantillon PPS de a UPE et d'un échantillon *epsem* de b ménages de chaque UPE sélectionnée produit un modèle global *epsem*. Avec un tel modèle, l'équation (18) reste valable, mais avec ρ à présent interprété comme une mesure synthétique de l'homogénéité à l'intérieur des grappes ultimes créées par le modèle de sous-échantillon (Kalton, 1979). La valeur de ρ , par exemple, pour un modèle de sous-échantillon qui sélectionne b ménages par échantillonnage systématique est différente de celle d'un modèle de sous-échantillon qui divise chaque UPE sélectionnée en sous-zones contenant b ménages chacune et sélectionne l'une de ces sous-zones (dans ce dernier cas, la valeur de ρ a des chances d'être plus élevée). Cette extrapolation traite donc à la fois de l'échantillonnage PPS et de diverses autres formes de conception de sous-échantillons.

31. Considérons à présent la stratification des UPE. Kalton (1979) montre que l'effet lié à la conception résultant de la mise en grappes d'un modèle global *epsem*, où un échantillon stratifié de a UPE est sélectionné et b unités élémentaires sont sélectionnées selon une probabilité égale à l'intérieur de chaque UPE sélectionnée, peut être donné de façon approximative :

$$D^2(\bar{y}_{cl}) = 1 + (b-1)\bar{\rho} \quad (19)$$

où $\bar{\rho}$ est la mesure moyenne de l'homogénéité à l'intérieur d'une strate, à condition que l'homogénéité à l'intérieur de chaque strate soit sensiblement de même grandeur. Le gain provenant de la stratification effective des UPE peut être substantiel lorsque b est de taille appréciable car la mesure globale de l'homogénéité donnée par l'équa-

³ L'analyse de la présente section s'applique aux mesures de l'homogénéité interne des grappes tant pour les grappes de taille égale qu'inégale.

tion (18) est remplacée par une mesure plus faible de l'homogénéité à l'intérieur d'une strate de l'équation (19). Exprimée autrement, la réduction de l'effet lié à la conception de $(b-1)(\rho - \bar{\rho})$ découlant de l'échantillonnage stratifié des UPE peut être plus forte lorsque b est élevé.

32. Jusqu'ici, nous avons supposé un échantillon *epsem* global où la taille de l'échantillon de chaque UPE sélectionnée est la même, b . Ces conditions sont réunies lorsque des UPE de taille égale sont sélectionnées selon une probabilité égale et lorsque des UPE de tailles inégales sont sélectionnées par échantillonnage PPS exact. Toutefois, dans la pratique, aucune de ces deux situations ne se produit. Des UPE de tailles assez inégales sont sélectionnées par PPES, avec des mesures estimatives de la taille qui sont sensiblement inexactes. En pareil cas, l'application de taux de sous-échantillonnage dans les UPE sélectionnées pour produire un modèle *epsem* global engendre une certaine variation de la taille des sous-échantillons. À condition que cette variation ne soit pas forte, l'équation (19) peut encore être utilisée comme approximation, b étant alors remplacé par la taille moyenne des sous-échantillons, c'est-à-dire :

$$D^2(\bar{y}_{cl}) = 1 + (\bar{b} - 1)\bar{\rho} \quad (20)$$

où $\bar{b} = \sum b_\alpha / a$ et b_α est le nombre d'unités élémentaires de l'UPE α . L'équation (20) s'est révélée d'une grande utilité pratique pour les situations où le nombre d'unités sélectionnées dans chaque UPE est relativement constant.

33. Toutefois, lorsque les différences de taille des sous-échantillons des UPE sont substantielles, l'approximation que comporte l'équation (20) devient inadéquate. Holt (1980) étend cette approximation en remplaçant \bar{b} dans l'équation (20) par une taille moyenne pondérée des sous-échantillons. L'effet lié à la conception résultant de la mise en grappes avec des grappes de tailles inégales peut être exprimé par l'équation :

$$D^2(\bar{y}_{cl}) = 1 + (b' - 1)\bar{\rho} \quad (21)$$

où $b' = \sum b_\alpha^2 / \sum b_\alpha$. (La grandeur b' peut être considérée comme la moyenne pondérée $b' = \sum k_\alpha b_\alpha / \sum k_\alpha$ où $k_\alpha = b_\alpha$) Comme précédemment, l'approximation suppose un modèle global d'échantillon *epsem*.

34. À titre d'exemple, supposons qu'il y ait cinq UPE sélectionnées avec des tailles de sous-échantillon de 10, 10, 20, 20 et 40 ménages, et supposons que $\bar{\rho} = 0,05$. La taille moyenne des sous-échantillons est $\bar{b} = 20$, tandis que $b' = 26$. Dans cet exemple, l'effet lié à la conception résultant de la mise en grappes est donc de 1,95 si l'on utilise l'approximation (20) et de 2,25 si l'on utilise l'approximation (21).

35. Verma, Scott et O'Muircheartaigh (1980) et Verma et Lê (1996) proposent un autre moyen d'effectuer cet ajustement à utiliser lorsque les tailles des sous-échantillons sont très différentes selon les domaines (par exemple, les domaines urbains et ruraux). Avec deux domaines, supposons que b_1 ménages sont sélectionnés dans chacune des a_1 UPE sélectionnées dans un domaine, avec $n_1 = a_1 b_1$, et que b_2 ménages sont sélectionnés dans le reste a_2 des UPE sélectionnées dans l'autre domaine, avec $n_2 = a_2 b_2$. Dès lors :

$$b' = (n_1 b_1 + n_2 b_2) / (n_1 + n_2)$$

36. Dans l'analyse précédente, on a considéré les effets liés à la conception résultant de la mise en grappes pour des estimations de la moyenne (et des proportions) de la population totale. Ce raisonnement est en grande partie applicable aux estimations des sous-groupes, à condition que l'on tienne dûment compte des hypothèses sous-jacentes. Il est utile d'introduire une triple classification des sous-groupes, selon leur distribution à travers les UPE. À une extrémité, il y a les sous-groupes qui sont répartis de façon régulière entre les UPE et qui sont appelés les « catégories croisées ».

Par exemple, les sous-groupes d'âge et de sexe sont généralement des catégories croisées. À l'autre extrémité, il y a les sous-groupes dont chacun est concentré sur un sous-ensemble d'UPE, et qui sont appelés « catégories différenciées ». Les sous-groupes urbains et ruraux sont généralement de ce type. Entre les deux se trouvent les sous-groupes qui sont plus ou moins concentrés dans certaines UPE. Il s'agit des « catégories mixtes ».

37. Les catégories croisées suivent la distribution de l'échantillon total à travers les UPE. Si l'échantillon total est assez régulièrement réparti entre les UPE, on peut alors utiliser l'équation (20) pour calculer la valeur approximative de l'effet lié à la conception résultant de la mise en grappes; on peut également utiliser cette équation pour une catégorie croisée, mais dans ce cas, il se produit un important changement : \bar{b} représente alors la taille moyenne du sous-échantillon de la catégorie croisée de chaque UPE. Du fait de ce changement, les effets liés à la conception pour les estimations de catégorie croisée sont moins prononcés que pour les estimations de l'échantillon total.

38. Les catégories séparées représentent toutes les unités d'un sous-ensemble d'UPE de l'échantillon complet. Comme la taille de l'échantillon d'une sous-catégorie pour une catégorie différenciée est la même que celle de l'échantillon total de ce sous-ensemble d'UPE, en général, il n'y a pas de raison d'attendre que l'effet lié à la conception pour une estimation correspondant à la catégorie différenciée soit moins prononcé que celui d'une estimation de l'échantillon total. L'effet lié à la conception pour une estimation portant sur une catégorie différenciée ne diffère de celui correspondant à une estimation de l'échantillon total que si la taille moyenne du sous-échantillon par UPE de la catégorie différenciée diffère de celle observée dans l'échantillon total ou si l'homogénéité est différente (et comprend, par exemple, une différence de la valeur synthétique ρ due à des conceptions différentes des sous-échantillons de la catégorie différenciée et d'ailleurs). Si l'échantillon total est réparti de façon régulière à travers les UPE, on peut à nouveau appliquer l'équation (20), avec \bar{b} et ρ comme valeurs pour l'ensemble d'UPE de la catégorie différenciée.

39. Lorsque la distribution d'une catégorie mixte est inégale à travers les UPE, cela implique que l'équation (20) n'est pas applicable. Pour estimer l'effet lié à la conception résultant de la mise en grappes pour une estimation faite à partir d'une catégorie mixte, on peut utiliser l'équation (21) avec b_α indiquant le nombre de membres sélectionnés d'une catégorie mixte dans l'UPE α .

3. Ajustements des pondérations

40. Comme on l'a vu à la section B.1, instituée « Stratification », les probabilités de sélection inégales entre des strates à stratification disproportionnée obligent à utiliser des pondérations dans l'analyse des données de l'enquête. Les équations (15) et (16) donnent l'effet lié à la conception résultant de la stratification disproportionnée et produisant des pondérations inégales dans l'hypothèse où la moyenne des strates et les variances unitaires sont toutes égales. Nous examinerons à présent d'autres variantes de ces formules qui se prêtent mieux à la détermination des effets des pondérations au stade de l'analyse. Premièrement, toutefois, nous prenons acte des facteurs qui donnent lieu à la nécessité de pondérations variables dans l'analyse de l'enquête [voir également Kish (1992)]. En premier lieu, comme nous l'avons déjà souligné, il faut des pondérations variables dans l'analyse pour compenser l'inégalité des probabilités de sélection liée à la stratification disproportionnée. D'une façon plus générale, ces pondérations sont nécessaires pour compenser l'inégalité des probabilités de sélection découlant d'une cause quelconque. Les pondérations qui compensent l'inéga-

lité des probabilités de sélection sont l'inverse de ces probabilités, et elles sont souvent appelées pondérations de base. Ces pondérations de base sont souvent ajustées pour compenser les cas de non-réponse et pour faire en sorte que les totaux d'échantillons pondérés soient conformes aux totaux connus de la population. De ce fait, les pondérations finales sont presque toujours variables dans une certaine mesure.

41. Même sans suréchantillonnage de certains domaines, les modèles d'échantillons s'écartent généralement du modèle *epsem* en raison de problèmes de cadre. Par exemple, si les ménages sont sélectionnés selon une probabilité égale à partir d'un cadre de ménages, puis un membre de chaque ménage est sélectionné au hasard, ces membres sont sélectionnés selon des probabilités inégales, et des pondérations sont nécessaires comme compensation pour l'analyse. Comme on le verra ci-dessous, ces pondérations engendrent un effet lié à la conception. On notera au passage que l'on peut éviter cet effet de la pondération en prenant tous les membres des ménages sélectionnés pour former l'échantillon. Toutefois, cette procédure introduit une nouvelle phase de mise en grappes, avec un nouvel effet de mise en grappes dû à la similarité de nombreuses caractéristiques des membres des ménages [voir Clark et Steel (2002) sur les effets liés à la conception que comportent ces différentes méthodes de sélection de personnes dans les ménages sélectionnés].

42. L'autre exemple courant de modèle non-*epsem* résultant d'un problème de cadre est celui où un modèle d'échantillon à deux phases est utilisé et où les unités primaires d'échantillonnage (UPE) sont sélectionnées selon des probabilités proportionnelles aux tailles estimatives (PPES). Si la mesure de la taille est raisonnablement exacte, la taille de l'échantillon par UPE sélectionnée pour un modèle global *epsem* est sensiblement la même pour toutes les UPE. En revanche, si la taille estimative d'une UPE sélectionnée représente une sérieuse sous-estimation, le modèle *epsem* demande un nombre d'unités de beaucoup supérieur à la moyenne pour cette UPE. Comme la collecte de données pour un aussi grand nombre est souvent impossible, un échantillon plus petit peut être prélevé, ce qui entraîne une sélection inégale des probabilités et la nécessité de pondérations compensatoires.

43. Pratiquement toutes les enquêtes se heurtent au même niveau de non-réponse. Une méthode couramment utilisée pour réduire la distorsion causée par la non-réponse consiste à appliquer un ajustement différentiel aux pondérations de base des déclarants. Cette procédure consiste à identifier des sous-groupes de l'échantillon qui affichent des taux de réponse différents et à gonfler les pondérations des déclarants de chaque sous-groupe en leur appliquant l'inverse du taux de réponse de ce sous-groupe (Brick et Kalton, 1996). Ces ajustements de pondération entraînent des écarts entre les pondérations et les pondérations de base qui ont souvent pour effet d'accroître l'effet lié à la conception d'une estimation.

44. Lorsque l'on peut obtenir les informations recherchées sur la population d'une autre source, les pondérations ajustées pour tenir compte du taux de non-réponse peuvent subir un nouvel ajustement pour que les estimations pondérées de l'échantillon soient conformes aux informations sur la population en question. Par exemple, si l'on peut obtenir d'une source extérieure de bonnes estimations de la population d'une région, on peut faire coïncider les estimations de l'échantillon concernant cette population avec les estimations extérieures. Ce type d'ajustement de la pondération d'une population se fait souvent après stratification. Il peut aider à compenser la non-couverture et améliorer la précision de certaines estimations de l'enquête. En revanche, il ajoute encore à la variabilité des pondérations, ce qui peut nuire à la précision des estimations de l'enquête qui sont sans rapport avec les variables de population employées dans l'ajustement.

45. Dans ce contexte, nous examinerons à présent une généralisation de l'effet lié à la conception pour une stratification disproportionnée afin d'évaluer les effets généraux des pondérations des variables. Kish (1992) présente une autre façon d'exprimer l'effet lié à la conception pour une moyenne stratifiée qui est très utile pour le calcul de l'effet d'une stratification disproportionnée au stade de l'analyse. L'équation suivante est simplement une forme différente des équations (15) et (16), et repose donc sur les mêmes hypothèses de moyennes des strates et surtout de variances unitaires égales. Comme il se calcule à partir de l'échantillon, l'effet lié à la conception est désigné comme $d^2(\bar{y}_{st})$ et

$$d^2(\bar{y}_{st}) = \frac{n \sum_h \sum_i w_{hi}^2}{(\sum_h \sum_i w_{hi})^2} = 1 + cv^2(w_{hi}) \quad (22)$$

où $cv(w_{hi})$ est le coefficient de variation des pondérations,

$cv^2(w_{hi}) = \sum \sum (w_{hi} - \bar{w})^2 / n\bar{w}^2$, et $\bar{w} = \sum \sum w_{hi} / n$ est la moyenne des pondérations.

46. Cette équation peut être présentée sous une forme plus générale :

$$d^2(\bar{y}_{st}) = \frac{n \sum_j w_j^2}{(\sum_j w_j)^2} = 1 + cv^2(w_j) \quad (23)$$

où chacune des n unités de l'échantillon a sa propre pondération w_j ($j = 1, 2, \dots, n$). L'effet lié à la conception résultant de pondérations inégales, exprimé par l'équation (23) est fonction de l'hypothèse selon laquelle les pondérations sont sans rapport avec la variable de l'enquête. L'équation peut donner une mesure raisonnable de l'effet d'une pondération différentielle pour des probabilités de sélection inégales si les hypothèses sur lesquelles elle repose tiennent au moins approximativement [voir Spencer (2000), pour un effet approximatif lié à la conception pour le cas où les probabilités de sélection sont en corrélation avec la variable de l'enquête].

47. Les ajustements pour non-réponse se font généralement dans des catégories définies par des variables auxiliaires connues à la fois pour les déclarants et les non-déclarants. Pour réduire sensiblement la distorsion causée par les cas de non-réponse, les variables mesurées dans le cadre de l'enquête doivent nécessairement varier à travers ces catégories de pondérations. Toutefois, cette variation n'est généralement pas forte, surtout en ce qui concerne la variance unitaire. De ce fait, l'équation (23) est largement utilisée pour examiner l'effet des ajustements de pondération pour non-réponse sur la précision des estimations de l'enquête. On peut procéder à cet examen en calculant l'équation (23) uniquement avec les pondérations de base ou avec les pondérations des ajustements pour non-réponse. Si ce dernier calcul produit une valeur beaucoup plus grande que le premier, cela veut dire que les ajustements des pondérations pour non-réponse entraînent une forte perte de précision dans les estimations de l'enquête. Dans ce cas, il peut être recommandé de modifier les ajustements de pondération en faisant baisser considérablement les catégories de pondérations ou en allégeant à l'extrême les fortes pondérations afin de réduire la perte de précision.

48. Si l'équation (23) est valable pour la plupart des ajustements de pondération de l'échantillon pour non-réponse, elle donne rarement une bonne approximation de l'effet des ajustements de pondération de la population. En particulier, lorsque les pondérations sont post-stratifiées ou calibrées sur des totaux de contrôle connus d'une source extérieure, l'équation (23) donne une mauvaise approximation de l'effet lié à la conception pour la moyenne de y lorsque y est étroitement lié à un ou à plusieurs des totaux de contrôle. Par exemple, supposons que les pondérations sont post-stratifiées sur les totaux de contrôle de la population d'un pays par sexe. Considérons le cas extrême où les données de l'enquête sont utilisées pour estimer la proportion

des femmes dans la population. Dans ce cas de parfaite corrélation entre la variable y et la variable de contrôle, la proportion estimée n'est pas sujette à erreur d'échantillonnage et, par conséquent, a une variance zéro. Dans la pratique, la corrélation n'est pas parfaite, mais elle peut être considérable pour certaines des variables de l'enquête. Lorsque la corrélation est forte, la post-stratification ou la calibration sur des totaux de population connus peut améliorer sensiblement la précision des estimations de l'enquête, mais cette amélioration n'est pas visible lorsque l'on a utilisé l'équation (23). Au contraire, cette équation indique une perte de précision.

49. Il ressort de ce qui précède que l'équation (23) ne doit pas être utilisée pour estimer l'effet lié à la conception résultant des ajustements de pondération de la population pour des estimations fondées sur des variables qui sont étroitement liées aux variables de contrôle. Or, dans la plupart des enquêtes générales sur la population des pays en développement, il existe peu de variables de contrôle fiables, sinon aucune, et la relation entre toute variable de contrôle pouvant exister et les variables de l'enquête est rarement étroite. De ce fait, l'équation (23) ne devrait pas souvent causer un problème de forte surestimation des effets liés à la conception résultant de la pondération. Néanmoins, il faut savoir que l'équation (23) ne doit pas être appliquée aveuglément.

50. Nous concluons cette analyse des effets liés à la conception résultant de la pondération par quelques commentaires sur les effets de la pondération sur les estimations de sous-groupes. Tous les résultats présentés dans cette section et dans la section B.1 peuvent être appliqués sans réserve pour donner les effets liés à la conception pour les estimations de sous-groupes si on limite les calculs aux membres de ces sous-groupes. Toutefois, il faut faire preuve de prudence quand on essaie d'obtenir les effets liés à la conception résultant d'une pondération pour les estimations de sous-groupes à partir des résultats de l'échantillon complet. Pour qu'une telle déduction soit valable, il faut que la distribution des pondérations à l'intérieur du sous-groupe soit semblable à celle de l'échantillon complet. Parfois, tel est le cas, mais pas toujours. En particulier, lorsque l'on utilise une stratification disproportionnée pour donner des tailles adéquates aux échantillons de certains domaines (sous-groupes), les effets liés à la conception pour les estimations de l'échantillon total sont supérieurs à l'unité (dans les hypothèses de moyennes et de variances égales). Toutefois, les effets liés à la conception résultant d'estimations de domaines peuvent être égaux à 1 si l'on utilise d'égales probabilités de sélection à l'intérieur des domaines.

C. Modèles pour les effets liés à la conception

51. Dans la section précédente, nous avons présenté séparément certains résultats pour les effets liés à la conception résultant de la pondération et de la mise en grappes, en mettant principalement l'accent sur les effets liés à la conception pour les moyennes et les proportions. Dans la présente section, nous prolongerons ces résultats en considérant les effets liés à la conception résultant d'une combinaison de pondérations et de mises en grappes et les effets liés à la conception pour certains autres types d'estimations.

52. On a utilisé un certain nombre de modèles pour représenter les effets liés à la conception pour ces extensions. Ces modèles ont été utilisés tant pour la conception que pour l'analyse d'échantillons de conception complexe (Kalton, 1977; Wolter, 1985). Traditionnellement, ces modèles ont joué un rôle important dans l'analyse. Toutefois, leur utilisation pour cet usage est en train de disparaître. Leur principale utilisation à venir, importante de surcroît, dans la planification de nouveaux modèles sera l'objet de la présente discussion.

53. Ces dernières années, des progrès considérables ont été réalisés dans les domaines de l'informatique et des logiciels de calcul des erreurs provenant de la complexité de conception des échantillons. Auparavant, le calcul des erreurs d'échantillonnage sur les estimations faites à partir d'échantillons complexes était une tâche laborieuse. Il était donc courant de calculer ces erreurs directement à partir d'un nombre d'estimations relativement faible et d'utiliser l'effet lié à la conception ou d'autres modèles pour déduire les erreurs d'échantillonnage commises sur d'autres estimations. Aujourd'hui, la situation s'est profondément améliorée, et le calcul direct des erreurs d'échantillonnage de nombreuses estimations ne pose plus de problème majeur. De plus, on peut compter sur de nouveaux progrès à venir tant en ce qui concerne la puissance des ordinateurs que les logiciels. Par conséquent, l'utilisation pour ce calcul des modèles d'effets liés à la conception est essentiellement vouée à disparaître.

54. L'utilisation du calcul des erreurs d'échantillonnage au stade de l'analyse a également pour but d'offrir un moyen de résumer succinctement ces erreurs dans les rapports d'enquête, et d'éliminer ainsi la nécessité de présenter individuellement une erreur d'échantillonnage pour chaque estimation. Dans certains cas, on peut également faire valoir que les estimations d'erreurs d'échantillonnage à partir d'un modèle peuvent être préférables à des estimations directes de ces erreurs car elles sont plus précises. Parfois, cet argument a une certaine force (par exemple, quand il s'agit d'estimer l'erreur d'échantillonnage sur une estimation effectuée dans une région où le nombre d'UPE sélectionnées est très faible). Néanmoins, en règle générale, quelle que soit la raison, l'utilisation de modèles pour calculer les erreurs d'échantillonnage est discutable. La validité des estimations d'un modèle dépend de la validité du modèle lui-même et, lorsque l'on compare les erreurs d'échantillonnages obtenues par calcul direct et par estimation à partir d'un modèle, ces comparaisons mettent sérieusement en doute la validité des modèles [voir, par exemple, Bye et Gallicchio (1989)]. Par ailleurs, si les modèles utilisés pour estimer les erreurs d'échantillonnage offrent un moyen de résumer succinctement ces erreurs dans les rapports d'enquête, ils imposent aux usagers la charge d'avoir à effectuer les calculs de ces erreurs à partir de ces modèles. Notre conclusion générale est que les modèles fondés sur l'effet lié à la conception et les autres modèles utilisés pour estimer les erreurs d'échantillonnage joueront vraisemblablement à l'avenir un rôle limité dans l'analyse des résultats des enquêtes.

55. En revanche, les modèles fondés sur l'effet lié à la conception continueront de jouer un rôle important dans la conception des échantillons. Une bonne compréhension des conséquences d'une allocation disproportionnée de l'échantillon et des effets de la mise en grappes sur la précision des différents modèles d'estimation d'enquêtes est essentielle à une bonne conception des échantillons. Bien entendu, la détermination de la taille à donner à l'échantillon pour assurer une précision adéquate des principales estimations de l'enquête nécessite que l'on tienne compte de l'effet lié à la conception résultant d'une conception donnée. En outre, on peut améliorer la structure d'un modèle d'échantillon efficace en examinant les résultats de modèles résultant de différents types de conception. Il convient de noter que l'on a besoin d'estimations des paramètres inconnus, tels que ρ , pour appliquer les modèles au stade de la conception. On a donc besoin d'établir des estimations de ces paramètres à partir d'enquêtes passées, comme on le verra dans la section suivante.

56. Nous commencerons par décrire les modèles utilisés pour déterminer les effets de la mise en grappes dans les échantillons *epsem* sur un ensemble de statistiques allant au-delà des moyennes et des proportions examinées à la section B.3 intitulée « Ajustements des pondérations ». Pour présenter ces modèles, nous retournerons aux moyennes des sous-groupes déjà traitées, en faisant une distinction entre les catégories croisées, les catégories différenciées et les catégories mixtes. Pour une catégorie croisée

d , répartie de façon régulière à travers les UPE, l'effet lié à la conception est donné approximativement par l'équation (20) que nous reproduirons ici comme :

$$D^2(\bar{y}_{ct:d}) = 1 + (\bar{b}_d - 1)\bar{\rho}_d \quad (24)$$

où \bar{b}_d indique la taille moyenne de l'échantillon de la catégorie croisée par UPE et $\bar{\rho}_d$ est la mesure synthétique de l'homogénéité de y dans les UPE pour la catégorie croisée. Un modèle amplement utilisé suppose que la mesure de l'homogénéité pour la catégorie croisée est la même que celle de l'ensemble de la population, autrement dit, que $\bar{\rho}_d = \bar{\rho}$. Dès lors, l'effet lié à la conception pour la moyenne de la catégorie croisée peut être estimé par l'équation :

$$d^2(\bar{y}_{ct:d}) = 1 + (\bar{b}_d - 1)\hat{\rho} \quad (25)$$

où $\hat{\rho}$ est une estimation de $\bar{\rho}$ de l'échantillon complet donné par :

$$\hat{\rho} = \frac{d^2(\bar{y}_{cl}) - 1}{\bar{b} - 1} \quad (26)$$

57. On étend souvent l'application de cette approche au calcul de $\hat{\rho}$ pour un ensemble d'estimations comparables faisant intervenir des variables apparentées et, à condition que les valeurs de $\hat{\rho}$ soient relativement semblables, à l'utilisation d'une certaine forme de moyenne de ces valeurs pour estimer $\bar{\rho}$ et par conséquent également $\bar{\rho}_d$ pour des estimations de sous-groupes pour toutes les variables. Cette approche est souvent utilisée pour produire des modèles de l'effet lié à la conception afin de récapituler les erreurs d'échantillonnage dans les rapports d'enquête. Elle sert également de base à une forme de fonction généralisée de variance [*generalized variance function* (GVF)] utilisée à cette fin (Wolter, 1985, p. 204).

58. On se trouve en présence d'un cas particulier d'application de cette approche avec les estimations d'enquêtes qui sont des proportions de sous-groupes entrant dans différentes catégories d'une variable, telles que les proportions de différents sous-groupes qui ont atteint des niveaux d'instruction différents ou qui entrent dans différentes catégories professionnelles. On suppose souvent que la valeur de $\bar{\rho}$ ne doit être estimée que pour un classement par catégories, et qu'une fois estimé, $\hat{\rho}$ peut alors être appliqué à tous les classements par catégories. L'hypothèse d'un $\bar{\rho}$ commun est mathématiquement correcte lorsqu'il n'y a que deux catégories (par exemple, les ménages avec et les ménages sans l'électricité), mais elle ne tient pas nécessairement lorsqu'il y a plus de deux catégories. Considérons, par exemple, les estimations des travailleurs travaillant dans l'agriculture et dans les mines. La valeur de $\bar{\rho}$ pour les travailleurs agricoles est presque certainement beaucoup plus faible que celle des mineurs, car il y a des chances que le secteur minier soit concentré dans quelques zones. L'hypothèse d'une valeur commune de $\bar{\rho}$ pour toutes les catégories ne saurait donc être appliquée inconsidérément.

59. Lorsque les variances d'une moyenne par catégorie croisée obtenues à l'aide de l'équation (25) ont été comparées à celles calculées directement, on a constaté qu'elles avaient tendance à être sous-estimées. Cela peut être dû au fait que, bien que classés comme catégories croisées, les sous-groupes ne sont pas répartis de façon totalement régulière à travers les UPE. La solution qui a été employée pour remédier à ce problème a été de modifier l'équation (25), de sorte que :

$$d^2(\bar{y}_{ct:d}) = 1 + k_d(\bar{b}_d - 1)\hat{\rho} \quad (27)$$

où $k_d > 1$. Fondant ses travaux sur de nombreuses analyses empiriques, Kish (1995) suggère des valeurs de $k_d = 1,2$ ou $1,3$; Verma et Lê (1996) laissent k_d varier selon la taille de

la catégorie croisée (k_d étant toujours supérieur à 1). L'autre solution possible serait de remplacer \bar{b}_d dans (25) par $b'_d = \sum b_{d\alpha}^2 / \sum b_{d\alpha}$ conformément à l'équation (21).

60. Considérons à présent brièvement les effets liés à la conception pour les statistiques analytiques. La forme la plus simple et la plus largement utilisée de statistique analytique est la différence entre la moyenne ou la proportion de deux sous-groupes. On a généralement constaté que l'effet lié à la conception pour la différence entre deux moyennes est supérieur à 1, mais qu'il est inférieur à celui obtenu quand on traite les moyennes des deux sous-groupes comme des grandeurs indépendantes (Kish et Frankel, 1974; Kish, 1995). Exprimé en termes de variances :

$$V(\bar{y}_{u;d}) + V(\bar{y}_{u;d'}) < V(\bar{y}_{cl;d} - \bar{y}_{cl;d'}) < V(\bar{y}_{cl;d}) + V(\bar{y}_{cl;d'}) \quad (28)$$

où d et d' représentent les deux sous-groupes. La variance de la différence dans la moyenne est généralement inférieure à la limite supérieure lorsque les sous-groupes sont tous deux représentés dans les mêmes UPE. Cela se traduit par une covariance entre les deux moyennes qui est presque toujours positive, et cette covariance positive diminue ensuite la variance de la différence. Cet effet ne se produit pas lorsque les sous-groupes sont des catégories différenciées qui sont des ensembles différents d'UPE : dans ce cas, la limite supérieure s'applique. Dans l'hypothèse où les variances unitaires des deux sous-groupes sont les mêmes (autrement dit, où $S_d^2 = S_{d'}^2$), cette inégalité se ramène à :

$$1 < D^2(\bar{y}_d - \bar{y}_{d'}) < \frac{n_{d'} D^2(\bar{y}_d) + n_d D^2(\bar{y}_{d'})}{n_d + n_{d'}}$$

61. On est en présence d'un cas particulier de différence entre deux proportions lorsque les proportions sont toutes deux fondées sur la même variable de catégories multiples, comme c'est le cas, par exemple, lorsqu'il est demandé aux déclarants de faire un choix entre plusieurs options et que l'analyste cherche à déterminer si une option est plus populaire qu'une autre. Kish *et al.* (1995) ont examiné les effets liés à la conception pour ces différences et constaté empiriquement que $d^2(\rho_d - \rho_{d'}) = [d^2(\rho_d) + d^2(\rho_{d'})] / 4$ dans ce cas particulier.

62. La constatation donnée ci-dessus que les effets liés à la conception résultant de la mise en grappes sont généralement plus faibles pour les différences de moyenne que pour la moyenne globale s'applique aux autres statistiques analytiques. Voir Kish et Frankel (1974) pour quelques constatations empiriques et quelques suggestions concernant les effets liés à la conception pour les coefficients de régression multiples. Les effets liés à la conception pour les coefficients de régression sont les mêmes que pour les différences entre moyennes. Le fait que cela est conforme aux attentes apparaît lorsque l'on note que la pente d'une régression linéaire simple de y sur x peut être estimée de façon relativement efficace par $b = (\bar{y}_u - \bar{y}_l) / (\bar{x}_u - \bar{x}_l)$, où les moyennes de y et de x sont calculées pour les tiers supérieur (u) et inférieur (l) de l'échantillon fondé sur la variable x . Voir Skinner, Holt et Smith (1989) et Lehtonen et Pahkinen (1994) pour les effets liés à la conception en régression et d'autres formes d'analyse, et Korn et Graubard (1999) pour les effets liés à la conception d'échantillons complexes sur la précision de l'analyse des données de l'enquête.

63. Nous concluons cette section par quelques commentaires sur le problème éprouvant de la décomposition de l'effet global lié à la conception en composantes en raison de la pondération et de la mise en grappes. Le calcul de l'effet lié à la conception $d^2(\bar{y}) = v_c(\bar{y}) / v_u(\bar{y})$ porte sur les effets combinés de la pondération et de la mise en grappes. Toutefois, lorsque l'on utilisera les données de l'enquête actuelle pour pré-

parer une enquête à venir, il faudra séparer ces deux composantes de l'effet lié à la conception. Par exemple, la prochaine enquête peut être envisagée comme une enquête utilisant *epsem* alors que l'enquête en cours peut avoir suréchantillonné certains domaines. En outre, même si elle utilise les mêmes UPE et la stratification, la prochaine enquête choisira peut-être de modifier la taille des sous-échantillons par UPE. Kish (1995) examine cette question, pour laquelle il n'y a pas de solution unique ou simple. Ici, nous proposons une approche que ne peut être utilisée que lorsque les pondérations sont aléatoires ou plus ou moins aléatoires. Dans ce cas, l'effet global lié à la conception peut être décomposé approximativement en un produit des effets liés à la conception résultant de la pondération et de la mise en grappes, à savoir :

$$d^2(\bar{y}) = d_w^2(\bar{y}) \cdot d_{ci}^2(\bar{y}) \quad (29)$$

où $d_w^2(\bar{y})$ est l'effet lié à la conception résultant de la pondération, donné par l'équation (23) et $d_{ci}^2(\bar{y})$ est l'effet lié à la conception résultant de la mise en grappes donné par les équations (20) ou (21). Il y a peu de justification théorique à l'utilisation de l'équation (29); toutefois, utilisant une approche de modélisation, Gabler, Haeder et Lahiri (1999) déterminent l'effet lié à la conception donné par l'équation (29) comme limite supérieure. En utilisant l'équation (29) avec l'équation (20), on estime donc $\bar{\rho}$ par l'équation :

$$\hat{\bar{\rho}} = \frac{[d^2(\bar{y})/d_w^2(\bar{y})] - 1}{\bar{b} - 1} \quad (30)$$

Comme on le verra ci-dessous, aux fins de planification, l'estimation du paramètre $\bar{\rho}$ est plus importante que celle de l'effet lié à la conception résultant de la mise en grappes car ce paramètre est plus facilement utilisé pour différents types de conceptions. L'effet lié à la conception résultant de la mise en grappes dans une enquête ne peut être appliqué directement à la planification d'une autre enquête que si la taille des sous-échantillons par UPE reste la même.

D. Utilisation des effets liés à la conception pour la conception des échantillons

64. Les modèles examinés jusqu'ici dans ce chapitre à propos des effets liés à la conception peuvent aider à planifier la conception d'un nouvel échantillon. Toutefois, ils ont besoin d'être appuyés par des données empiriques, notamment pour la mesure synthétique de l'homogénéité $\bar{\rho}$. On peut obtenir ces données en analysant les effets liés à la conception sur des enquêtes passées de type similaire. L'accumulation de données sur ces effets est donc utile.

65. On dispose d'une quantité substantielle de données sur les effets liés à la conception pour les enquêtes démographiques de fécondité et de santé, qui proviennent d'analyses détaillées des erreurs d'échantillonnage effectuées pour les programmes d'Enquêtes mondiales de fécondité [*World Fertility Surveys* (WFS)] et d'Enquêtes démographiques et sanitaires. [*Demographic and Health Surveys* (DHS)]. Le programme WFS a porté sur 42 enquêtes effectuées dans 41 pays entre 1974 et 1982. Le programme DHS a suivi en 1984, avec des enquêtes effectuées dans la plupart des pays tous les trois à cinq ans. Voir Verma et Lê (1996) pour les analyses des erreurs d'échantillonnage commises dans le cadre des enquêtes DHS, Kish, Groves et Krotki (1976) et Verma, Scott et O'Muircheartaigh (1980) pour des analyses similaires des erreurs d'échantillonnage du programme WFS. Ces analyses ont révélé que les estimations de $\bar{\rho}$ sont relativement applicables à tous les pays, à condition que les échantillons soient de conception comparable. Autrement dit, quand on conçoit une nouvelle enquête

dans un pays donné, les données empiriques sur les erreurs d'échantillonnage constatées dans une enquête similaire effectuée dans un pays voisin peuvent être appliquées s'il y a lieu et si l'on prend soin de s'assurer que les échantillons sont de conception comparable.

66. L'exemple donné ci-dessous illustre l'utilisation des effets liés à la conception pour la conception d'un échantillon pour une enquête nationale hypothétique. Pour les besoins de cette illustration, nous supposons que l'échantillon est un échantillon stratifié à deux phases PPS, par exemple, avec les districts d'énumération du recensement comme UPE et les ménages comme unités de deuxième phase. Nous supposons que la principale statistique étudiée est la proportion de ménages vivant en dessous du seuil de pauvreté que, pour les besoins de la planification, on suppose être d'environ 25 %, et similaire pour toutes les provinces de ce pays. Les spécifications initiales sont que l'estimation de cette proportion devrait avoir un coefficient de variation de pas plus de 5 % pour le pays et de pas plus de 10 % pour chacune de ses huit provinces. En outre, l'échantillon devrait faciliter la production d'estimations précises d'un éventail de statistiques pour certains sous-groupes nationaux qui sont répartis assez régulièrement à travers ces huit provinces. Si l'on utilisait l'échantillonnage aléatoire simple, le coefficient de variation serait :

$$CV = \sqrt{\frac{1-P}{nP}}$$

où P est la proportion de ménages vivant en dessous du seuil de pauvreté (en l'occurrence, 25 %). Cette formule peut aussi être utilisée avec un échantillon complexe, à condition que n soit remplacé par la taille effective de l'échantillon, $n_{eff} = n/D^2(p)$.

67. La première question qui se pose est de savoir comment l'échantillon devrait être distribué à travers les provinces. Le tableau VI.2 indique la distribution de la population à travers les provinces (W_h), avec une allocation proportionnée de l'échantillon à travers les provinces, une allocation égale de la taille de l'échantillon à chaque province et une allocation de compromis de l'échantillon se situant entre l'allocation proportionnée et l'allocation égale. On utilise ici une taille de l'échantillon total fixée arbitrairement à 5 000 ménages, qui peut être révisée par la suite, si besoin est.

Tableau VI.2
Distributions de la population et trois allocations possibles de l'échantillon à travers les huit provinces (A-H)

	A	B	C	D	E	F	G	H	Total
W_h	0,33	0,24	0,20	0,10	0,05	0,04	0,02	0,02	1,00
Allocation proportionnée	1 650	1 200	1 000	500	250	200	100	100	5 000
Allocation à tailles égales de l'échantillon	625	625	625	625	625	625	625	625	5 000
Allocation de compromis de l'échantillon	1 147	879	767	520	438	427	411	411	5 000

68. Toutes choses égales par ailleurs, l'allocation proportionnée est celle qui convient le mieux pour produire des estimations nationales et des estimations pour des sous-groupes qui sont répartis régulièrement entre toutes les provinces. En revanche, l'allocation de l'échantillon à tailles égales est celle qui convient le mieux pour les estimations provinciales. Comme le montre le tableau VI.2, ces deux allocations diffèrent sensiblement, en raison des tailles très différentes des provinces indiquées à la

rangée W_h . L'allocation proportionnée donne pour les petites provinces (E, F, G et H) des échantillons qui sont trop petits pour permettre d'y calculer des estimations fiables. Inversement, l'allocation à tailles égales de l'échantillon diminue la précision des estimations nationales. Cette baisse de précision peut se calculer à l'aide de l'équation (15), qui, dans ce cas, se résume à $H \sum W_h^2 = 1,77$, où H est le nombre de provinces. Ainsi, si l'on ne considère que les effets de l'allocation disproportionnée (c'est-à-dire si l'on exclut les effets de la mise en grappes), la taille de 5 000 ménages de l'échantillon national est réduite à une taille effective de $5\,000/1,77 = 2\,825$.

69. La réponse à la question de savoir si la forte perte de précision dans les estimations nationales (en particulier pour les sous-groupes) par suite de l'utilisation de l'allocation à égalité est acceptable dépend de l'importance relative des estimations nationales et provinciales. Souvent, les estimations nationales sont suffisamment importantes pour que cette perte soit inacceptable. L'allocation de compromis de la dernière rangée du tableau VI.2 se calcule selon une allocation proposée par Kish (1976, 1988) pour le cas où les estimations nationales et provinciales sont d'égale importance. Cette allocation, donnée par $n_h \propto \sqrt{W_h^2 + H^{-2}}$, augmente considérablement la taille des échantillons pour les petites provinces par comparaison avec l'allocation proportionnée, mais pas autant que l'allocation à égalité. L'effet lié à la conception d'une pondération inégale pour cette allocation est de 1,22, contre 1,77 pour l'allocation de la taille de l'échantillon à égalité. Nous supposons que cette allocation de compromis est celle qui est adoptée pour l'enquête.

70. La question suivante qui se pose est de savoir comment déterminer le nombre d'UPE et le nombre désiré de ménages à choisir par UPE. Comme on l'a vu au chapitre II, si l'on utilise un modèle de coût simple, le nombre optimum de ménages à sélectionner par UPE de l'échantillon est donné par :

$$b_{opt} = \sqrt{C^* \frac{(1-\rho)}{\rho}}$$

où C^* est le ratio du coût de l'addition d'une UPE à l'échantillon au coût de l'addition d'un ménage. Le modèle de coût est trop simplifié, et la formule pour b_{opt} ne doit pas être utilisée inconsidérément; cela dit, elle fournit encore des indications utiles.

71. Supposons que l'organisation du travail de terrain rende raisonnable l'utilisation du modèle de coût simple et qu'une analyse de la structure des coûts indique que C^* est environ de 16. Par ailleurs, supposons qu'une enquête précédente, ayant utilisé les mêmes UPE, ait produit une estimation de $\bar{\rho} = 0,05$ pour une caractéristique qui est étroitement liée à la pauvreté. En appliquant ces chiffres à la formule ci-dessus, on obtient $b_{opt} = 17,4$, chiffre qui, par simplification, est arrondi à 17. Souvent, dans la pratique, le ratio de coût C^* n'est pas constant à travers tout le pays; par exemple, ce ratio peut être beaucoup plus faible en milieu urbain qu'en milieu rural. Si tel est le cas, on peut utiliser différentes valeurs dans différentes parties du pays. Une telle complexité n'est pas retenue ici. On trouvera des exemples de telles différences dans plusieurs des chapitres de cette publication qui décrivent des modèles d'échantillon national.

72. Avec $\bar{\rho} = 0,05$ et $b = 17$, l'effet lié à la conception résultant de la mise en grappes est :

$$D^2(p) = 1 + (b-1)\bar{\rho} = 1,80$$

Il faut tenir compte de cet effet pour la détermination de la précision des estimations provinciales. Par exemple, la taille effective de l'échantillon de 411 ménages de la province H est $411/1,80 = 228$. Par conséquent, le coefficient de variation pour la propor-

tion de ménages vivant en dessous du seuil de pauvreté dans la province H est de 0,11. Si ce niveau de précision était jugé insuffisant, il faudrait accroître la taille de l'échantillon dans la province H (ainsi que dans la province G).

73. Pour les estimations nationales, l'effet lié à la conception a besoin d'être combiné aux effets liés à la conception résultant de la mise en grappes et de l'allocation disproportionnée à travers les provinces. Ainsi, pour la proportion nationale de ménages vivant en dessous du seuil de pauvreté, on peut obtenir l'estimation de l'effet lié à la conception avec l'équation (29) : $1,22 \times 1,80 = 2,20$. Autrement dit, la taille effective de l'échantillon correspondant à un échantillon de 5 000 ménages est de 2 277 et le coefficient de variation de l'estimation nationale de la proportion de ménages vivant en dessous du seuil de pauvreté est de 0,036. Souvent, la taille de l'échantillon global est plus que suffisante pour satisfaire aux exigences de précision des estimations de la population totale. Ce qui préoccupe davantage, ce sont les niveaux de précision pour les sous-groupes de population. Dans ce cas, l'effet lié à la conception résultant de la mise en grappes pour les catégories croisées réparties de façon régulière à travers les UPE est plus faible que celui correspondant à l'échantillon total, comme il est indiqué à la section C. Par exemple, considérons une catégorie croisée composée d'un tiers de la population. Dans ce cas, en appliquant la formule (27) avec $k_d = 1,2$ et $\bar{b}_d = 17/3$, on obtient un effet lié à la conception résultant de la mise en grappes de 1,23. En combinant cet effet à celui correspondant à une allocation disproportionnée à travers les provinces, on obtient un effet global lié à la conception pour l'estimation d'une catégorie croisée de $1,22 \times 1,23 = 1,50$, et une taille effective de l'échantillon de $5\,000 / (3 \times 1,50) = 1\,111$. Le coefficient estimatif de variation pour l'estimation d'une catégorie croisée est donc de 0,05.

74. On peut faire des calculs du type de ceux décrits ci-dessus pour évaluer la précision probable des estimations clés de l'enquête, et on peut modifier la taille des échantillons selon les besoins. Dans les estimations finales de la taille des échantillons, il faut tenir compte des cas de non-réponse. Par exemple, avec un taux de réponse assez uniforme de 90 % à travers tout le pays, il faut accroître la taille des échantillons calculée ci-dessus de 11 %. En outre, l'effet lié à la conception peut augmenter quelque peu par suite de la variation additionnelle des pondérations découlant des ajustements pour non-réponse. Lorsque l'on calcule les fractions d'échantillonnage à utiliser pour engendrer des échantillons des tailles désirées, il faut tenir compte des cas de non-couverture. Avec un taux de couverture de 90 %, il faut accroître les fractions d'échantillonnage de 11 %.

E. Conclusion

75. Il est important de comprendre les effets liés à la conception et leurs composantes lorsque l'on cherche à concevoir des échantillons pour de nouvelles enquêtes. Par exemple :

- L'ampleur des effets globaux liés à la conception pour certaines estimations peut être utilisée pour déterminer la taille à donner à l'échantillon. On peut calculer la taille à donner à l'échantillon pour produire le niveau de précision spécifié pour chaque estimation clé pour un échantillon sans restrictions, et multiplier ensuite la taille de cet échantillon par l'effet lié à la conception de l'estimation pour déterminer la taille à donner à l'échantillon pour cette estimation avec l'échantillon de conception complexe. On peut alors choisir la taille définitive de l'échantillon en examinant les tailles requises pour chacune des estimations (et prendre, peut-être, la plus grande de ces tailles).

- Lorsque l'on utilise un échantillon stratifié disproportionné pour déterminer des estimations de domaines présentant les niveaux de précision requis, on peut évaluer la perte de précision causée à l'estimation de l'échantillon total et aux sous-groupes qui couvrent tous les domaines en calculant l'effet lié à la conception dû à des pondérations variables. Si l'on constate que la perte est trop grande, il peut être indiqué d'apporter un changement aux conditions liées aux domaines, qui débouche sur des pondérations moins variables.
- Si l'effet lié à la conception résultant de la mise en grappes est très prononcé pour certaines estimations clés de l'enquête, il conviendrait d'envisager la possibilité d'accroître le nombre d'UPE sélectionnées (a) et d'utiliser des sous-échantillons de taille plus petite (b).

76. Si les formules présentées dans ce chapitre sont utiles à la conception des échantillons, elles ne sauraient être appliquées inconsidérément. Comme on l'a indiqué à plusieurs reprises, ces formules sont le produit d'un certain nombre d'hypothèses et de simplifications. Les usagers doivent se montrer sensibles à ces aspects et se poser la question de savoir si ces formules produiront les approximations raisonnables pour leur situation.

77. L'estimation des effets liés à la conception résultant de la mise en grappes nécessite des estimations de ρ valeurs pour les variables clés de l'enquête. Ces estimations sont inévitablement imparfaites, mais des estimations raisonnables peuvent suffire. En cherchant à utiliser une valeur de ρ plus élevée que prévu, on aboutit à la nécessité d'un échantillon de taille plus grande; cette stratégie est donc prudente.

78. Enfin, il convient de noter que le but de l'utilisation de ces modèles d'effets liés à la conception est de produire un échantillon efficace. L'incapacité des modèles à rester exacts entraîne une perte d'efficacité. En revanche, l'utilisation de modèles impropres pour concevoir l'échantillon n'influe pas sur la validité des estimations de l'enquête. Avec un échantillonnage de probabilités, les estimations de l'enquête restent des estimations valables des paramètres de la population.

RÉFÉRENCES

- Brick, J. M., et G. Kalton (1996). Handling missing data in survey research. *Statistical Methods in Medical Research*, vol. 5, pp. 215-238.
- Bye, B., et S. Gallicchio (1989). A note on sampling variance estimates for Social Security program participants from the Survey of Income and Program Participation. *United States Social Security Bulletin*, vol. 51, n° 10, pp. 4-21.
- Clark, R. G., et D.G. Steel (2002). The effect of using household as a sampling unit. *International Statistical Review*, vol. 70, pp. 289-314.
- Cochran, W. G. (1977). *Sampling Techniques*, 3^e éd. New York, Wiley.
- Gabler, S., S. Haeder et P. Lahiri (1999). A model based justification of Kish's formula for design effects for weighting and clustering. *Survey Methodology*, vol. 25, pp. 105-106.
- Holt, D. H. (1980). Discussion of the paper by Verma, V., C. Scott and C. O'Muircheartaigh: sample designs and sampling errors for the World Fertility Survey. *Journal of the Royal Statistical Society, Série A*, vol. 143, pp. 468-469.
- Kalton, G. (1977). Practical methods for estimating survey sampling errors. *Bulletin of the International Statistical Institute*, vol. 47, n° 3, pp. 495-514.
- _____ (1979). Ultimate cluster sampling. *Journal of the Royal Statistical Society, Série A*, vol. 142, pp. 210-222.
- Kish, L. (1965). *Survey Sampling*. New York, Wiley.

- _____ (1976). Optima and proxima in linear sample designs. *Journal of the Royal Statistical Society, Séries A*, vol. 139, pp. 80-95.
- _____ (1982). Design effect. Dans *Encyclopedia of Statistical Sciences*, vol. 2, S. Kotz et N. L. Johnson, eds., New York, Wiley, pp. 347-348.
- _____ (1988). Multi-purpose sample designs. *Survey Methodology*, vol. 14, pp. 19-32.
- _____ (1992). Weighting for unequal P_i . *Journal of Official Statistics*, vol. 8, pp. 183-200.
- _____ (1995). Methods for design effects. *Journal of Official Statistics*, vol. 11, pp. 55-77.
- _____, et M. R. Frankel (1974). Inference from complex samples. *Journal of the Royal Statistical Society, Séries B*, vol. 36, pp. 1-37.
- _____, et al. (1995). Design effects for correlated. *Survey Methodology*, vol. 21, pp. 117-124.
- _____, et al. (1976). *Sampling Errors in Fertility Surveys*. World Fertility Survey Occasional Paper, n° 17. La Haye, International Statistical Institute.
- Korn, E. L., et B. I. Graubard (1999). *Analysis of Health Surveys*. New York, Wiley.
- Lehtonen, R., et E. J. Pahkinen (1994). *Practical Methods for Design and Analysis of Complex Surveys*, édition révisée, Chichester, Royaume-Uni, Wiley.
- Lepkowski, J. M., et J. Bowles (1996). Sampling error software for personal computers. *Survey Statistician*, vol. 35, pp. 10-17.
- Nations Unies (1993). *National Household Survey Capability Programme: Sampling Errors in Household Surveys*. UNFPA/UN/INT-92-P80-15E. New York, Division de statistique des Nations Unies. Publication préparée par Vijay Verma.
- Rust, K. F. (1985). Variance estimation for complex estimators in sample surveys. *Journal of Official Statistics*, vol. 1, pp. 381-397.
- _____, et J. N. K. Rao (1996). Variance estimation for complex surveys using replication techniques. *Statistical Methods in Medical Research*, vol. 5, pp. 283-310.
- Skinner, C. J., D. Holt et T. M. F. Smith, eds. (1989). *Analysis of Complex Surveys*. Chichester, Royaume-Uni, Wiley.
- Spencer, B. D. (2000). An approximate design effect for unequal weighting when measurements may correlate with selection probabilities. *Survey Methodology*, vol. 26, pp. 137-138.
- Verma, V., et T. Lê (1996). An analysis of sampling errors for the Demographic and Health Surveys. *International Statistical Review*, vol. 64, pp. 265-294.
- Verma, V., C. Scott et C. O'Muircheartaigh (1980). Sample designs and sampling errors for the World Fertility Survey. *Journal of the Royal Statistical Society, Séries A*, vol. 143, pp. 431-473.
- Wolter, K. M. (1985). *Introduction to Variance Estimation*. New York, Springer-Verlag.

Chapitre VII

Analyse des effets liés à la conception pour les enquêtes dans les pays en développement

HANS PETTERSSON
Statistics Sweden
Stockholm, Suède

PEDRO LUIS DO NASCIMENTO SILVA
Escola Nacional de Ciências Estadísticas/
Instituto Brasileiro de Geografia e Estatística
(ENCE/IBGE)
Rio de Janeiro, Brésil

RÉSUMÉ

Le présent chapitre décrit les effets liés à la conception observés sur 11 enquêtes sur les ménages effectuées dans sept pays et sur trois enquêtes qui sont assez semblables par leur conception; compare les effets liés à la conception et les taux d'homogénéité (*roh*) pour des estimations de la consommation des ménages et de leur possession de biens durables. Il conclut par une discussion de la transférabilité des estimations de *roh* d'une enquête à une autre.

Termes clés : effets liés à la conception, efficacité, taux d'homogénéité, conception de l'enquête, conception d'échantillon, mise en grappes.

A. Introduction

1. Il n'est pas encore courant de calculer les effets liés à la conception comme résultat normal des enquêtes sur les ménages dans les pays en développement. Il y a toutefois des exceptions en ce qui concerne certaines enquêtes normalisées, comme celles du programme d'étude sur la mesure des niveaux de vie (LSMS) et les enquêtes démographiques et sanitaires (DHS), pour lesquelles les effets liés à la conception ont été calculés et comparés entre différents pays (voir chapitres XXII et XXIII). Une analyse comparative a déjà été faite sur 35 enquêtes effectuées dans le cadre du programme d'Enquêtes mondiales sur la fécondité (WFS) [Verma, Scott et O'Muircheartaigh, 1980].

2. Dans ce chapitre, nous présenterons les effets liés à la conception observés sur 11 enquêtes effectuées dans sept pays. Le choix de ces enquêtes était subjectif et fondé principalement sur la facilité d'y accéder. Les pays en question sont : le Brésil (3), le Cambodge (1), la République démocratique populaire lao (1), le Lesotho (1), la Namibie (2), l'Afrique du Sud (2) et le Viet Nam (1). Les enquêtes sont de nature différente et portent sur des sujets différents. Il s'agit notamment d'enquêtes polyvalentes, d'enquêtes sur la main-d'œuvre, d'une enquête sur les niveaux de vie et d'une enquête démographique. Les effets liés à la conception ont été calculés pour un certain nombre

de caractéristiques, principalement pour les besoins de la planification d'enquêtes. Ce chapitre a pour principal but de donner au lecteur une idée générale des niveaux des effets liés à la conception observés dans diverses enquêtes.

3. Pour trois enquêtes assez semblables dans leur conception, on a effectué une analyse plus approfondie et comparé les effets liés à la conception et les taux d'homogénéité de quelques variables concernant la consommation des ménages et l'accès aux biens durables. Il s'agissait d'examiner le comportement de la même variable (à peu de chose près) chez des populations différentes et d'étudier les similarités et les éventuelles tendances qui se dégagent des constatations faites.

Tableau VII.1

Caractéristiques des 11 enquêtes sur les ménages incluses dans cette étude

Enquêtes	Nombre de niveaux d'échantillonnage	Échantillon de première phase : nombre d'UPE sélectionnées pour l'échantillon	Taille des UPE : nombre moyen de ménages par UPE	Taille des grappes : Nombre de ménages sélectionnés par UPE (ou par USE, s'il y a deux niveaux d'échantillonnage)	Taille de l'échantillon : nombre de ménages de l'enquête	Allocation de l'échantillon entre les strates
Enquête sur les dépenses et la consommation en Rép. dém. pop. lao, 1997-1998	1	R : 348 U : 102	R : 51 U : 87	R : 20 U : 20	R : 6 960 U : 2 040	Disproportionnée
Enquête socio-économique au Cambodge, 1999	1	R : 360 U : 240	R : 154 U : 243	R : 10 U : 10	R : 3 600 U : 2 400	Approximativement proportionnée
Enquête sur le revenu et les dépenses des ménages en Namibie, 1993-1994	1	R : 123 U : 96	R : 152 U : 148	R : 20 U : 20	R : 2 685 U : 1 712	Approximativement proportionnée
Enquête démographique effectuée entre deux recensements en Namibie, 1995-1996	1	R : 120 U : 82	R : 152 U : 148	R : 50 U : 50	R : 5 600 U : 3 900	Approximativement proportionnée
Enquête polyvalente sur les ménages au Viet Nam, 1999	2	839 UPE, (2 USE choisies dans chaque UPE)	R : 1 417 U : 2 579 USES : R : 99 U : 105	R : 15 U : 15	25 170	Disproportionnée
Enquête sur la main-d'œuvre au Lesotho, 1997	1	R : 80 U : 40	R : 370 U : 341	R : 33 (moyenne) U : 25 (moyenne)	R : 2 600 U : 1 000	Approximativement proportionnée
Enquête de février 2000 sur la main-d'œuvre en République sud-africaine	1	R : 426 U : 1 148	R : min 100 ^a U : min 100 ^a	R : 10 U : 5	R : 4 059 U : 5 646	Disproportionnée
Enquête d'octobre 1999 sur les ménages en République sud-africaine	1	R : 1 273 U : 1 711	R : 110-120 U : 80-100	R : 10 U : 10	R : 10 923 U : 15 211	Disproportionnée
Enquête PNAD, 1999, Brésil	1 ou 2	7 019	250	13	93 959	Disproportionnée
Enquête PME de septembre 1999, Brésil	1	1 557	250	20	30 535	Disproportionnée
Enquête PPV, 1996-1997, Brésil	1	554	250	R : 16 U : 8	4 944	Fortement disproportionnée

Note : R = rural, U = urbain

^a Minimum de 100.

B. Les enquêtes

4. Les enquêtes pour lesquelles les effets liés à la conception sont indiqués dans ce chapitre sont les suivantes :

- L'Enquête sur les dépenses et la consommation en République démocratique populaire lao, 1997-1998 (LECS)
- L'Enquête socio-économique au Cambodge, 1999 (CSES)
- L'Enquête sur le revenu et les dépenses des ménages en Namibie, 1993-1994 (NHIES)

- L'Enquête démographique effectuée entre deux recensements en Namibie, 1995/96 (NIDS)
- L'Enquête polyvalente sur les ménages au Viet Nam, 1999 (VMPHS)
- L'Enquête sur la main-d'œuvre au Lesotho, 1997 (LFS)
- L'Enquête d'octobre 1999 sur les ménages en République sud-africaine (OHS)
- L'Enquête de février 2000 sur la main-d'œuvre en République sud-africaine
- PNAD (*Pesquisa Nacional por Amostra de Domicílios*) [Enquête nationale par échantillonnage sur les logements], 1999, Brésil
- PME (*Pesquisa Mensal de Emprego*) [Enquête mensuelle sur l'emploi] de septembre 1999, Brésil
- PPV (*Pesquisa de Padrões de Vida*) [Enquête sur les modes de vie], 1996-1997, Brésil

5. Le tableau VII.1 résume les principales caractéristiques de conception des 11 enquêtes. Celles-ci ont toutes utilisé des modèles standard à deux phases à probabilité proportionnelle à la taille (PPS), sauf l'enquête effectuée au Viet Nam, pour laquelle a été utilisée une conception en trois phases. La PNAD a également utilisé un échantillonnage en trois phases pour les petites municipalités non métropolitaines, mais celles-ci ne comptaient qu'environ un tiers de la population couverte par l'enquête. La plupart de ces enquêtes ont utilisé les zones d'énumération du recensement comme UPE (avec, dans certains cas, quelques modifications de petites ZE). Des tailles moyennes de 90 à 150 ménages par UPE étaient communes à ces enquêtes. Trois enquêtes se sont écartées de ce schéma. Les deux enquêtes du Lesotho avaient des UPE beaucoup plus grandes, composées de groupes de ZE comptant en moyenne de 340 à 370 ménages. À l'autre extrémité du spectre, les UPE rurales de la République démocratique populaire lao ne comptaient en moyenne que 50 ménages.

6. Pour plusieurs de ces enquêtes, la taille des échantillons prélevés dans les UPE (la taille des grappes) était d'environ 20 ménages. L'Enquête démographique effectuée entre deux recensements en Namibie se distingue par des échantillons de 50 ménages sélectionnés dans chaque UPE. À l'autre extrême, on trouve la PPV au Brésil, pour laquelle l'échantillon prélevé sur les UPE urbaines n'était que de 8 ménages, et les deux enquêtes effectuées en Afrique du Sud ainsi que celle effectuée au Cambodge, pour lesquelles il n'a été sélectionné que 10 ménages par UPE. La plupart de ces enquêtes avait la même taille de grappes entre zones urbaines et rurales.

7. La plupart de ces enquêtes étaient explicitement stratifiées en zones urbaines et rurales à l'intérieur des divisions administratives (provinces, régions). Au Lesotho, la LFS avait une stratification supplémentaire dans les zones agroécologiques, la LECS lao avait une stratification supplémentaire entre villages qui avaient ou non une route d'accès. Au Brésil, la PNAD et la PME n'étaient stratifiées qu'implicitement en zones rurales et urbaines, avec une sélection PPS systématique des UPE après classification par emplacement.

8. La sélection systématique a été utilisée pour la sélection des ménages à l'intérieur des unités géographiques ultimes pour toutes les enquêtes sauf l'enquête PPV, pour laquelle les ménages ont été sélectionnés par échantillonnage aléatoire simple.

9. L'un des traits importants de beaucoup de modèles d'échantillons est qu'ils ont utilisé des allocations disproportionnées dans toutes les provinces afin de produire des estimations provinciales d'une précision adéquate. Parfois, les pondérations nécessaires à l'analyse pour compenser ces allocations disproportionnées ont été très variables. Par exemple, pour la PPV brésilienne, le ratio de la plus forte pondération à

la plus faible était d'environ 40. On trouvera plus de détails à l'annexe sur les modèles d'échantillon utilisés pour les enquêtes.

C. Effets liés à la conception

10. Les effets liés à la conception [$d^2(\bar{y})$] sur une sélection d'estimations de chaque enquête sont indiqués aux tableaux VII.2 à VII.6 (pour une description de la façon dont est calculé l'effet lié à la conception, voir chapitre VI). Les effets liés à la conception ont été calculés à l'aide d'un Logiciel d'analyse statistique de données corrélées [*Software for the Statistical Analysis of Correlated Data* (SUDAAN) ou STATA]. Dans certains cas, ils ont été indiqués par l'office national de statistiques¹.

11. La variation de ces effets est substantielle, comme on pouvait s'y attendre, étant donné les différences dans la conception des échantillons et les variables utilisées par les enquêtes et les différences dues aux conditions propres à la population de chaque pays. Certains effets sont très sensibles. Des effets de l'ordre de 6 à 10 pour les variables sur les ménages ne sont pas inhabituels dans les résultats affichés aux tableaux VII.2 à VII.6, et certains effets se situent dans la fourchette de 10 à 15. Il convient de noter que ces effets liés à la conception reflètent les effets de conceptions complexes d'échantillons stratifiés en grappes et les allocations disproportionnées entre les provinces (le cas échéant). Les tableaux des effets liés à la conception présentés aux tableaux VII.2 à VII.6 donnent une illustration des niveaux de ces effets observés sur certaines enquêtes socio-économiques et démographiques sur les ménages dans les pays en développement.

12. Le tableau VII.2 présente des estimations des effets liés à la conception pour sept enquêtes effectuées en Afrique et en Asie du Sud-Est au niveau national et pour les domaines urbain et rural. La plupart de ces effets concernent les variables socio-économiques relatives à ces ménages. Les effets liés à la conception découlant de trois de ces enquêtes portent principalement sur les variables relatives à la main-d'œuvre à l'échelon individuel. Globalement, l'effet moyen lié à la conception à l'échelon national est de 4,2, mais on note des écarts substantiels de ces effets, qui varient de 1,3 à 8,1, la plupart d'entre eux se situant entre 2,0 et 6,0. Les effets moyens liés à la conception pour les sous-domaines urbain et rural sont respectivement de 4,1 et 4,0. Du fait des différences observées en ce qui concerne la conception des échantillons et les variables, il est difficile d'analyser les résultats du tableau pour en dégager des différences d'ordre général entre les types de variables (par exemple, socio-économiques ou main-d'œuvre) et de domaines (urbain/rural). On présentera au tableau VII.7 un essai de comparaison de certains des effets liés à la conception.

13. Le tableau VII.3 présente des estimations des effets liés à la conception pour un certain nombre d'estimations du niveau des ménages, tirées de l'enquête PNAD au Brésil.

¹ Le professeur David Stoker de Statistics South Africa a rassemblé des effets liés à la conception de l'Enquête sur la main-d'œuvre et de l'Enquête d'octobre sur les ménages de la République sud-africaine. Les effets liés à la conception de l'Enquête polyvalente sur les ménages du Viet Nam ont été communiqués par M. Nguyen Phong, Directeur du Département de statistiques sociales et environnementales de l'Office général de statistiques du Viet Nam. Les effets liés à la conception de l'Enquête sur le revenu et les dépenses des ménages en Namibie ont été calculés par M. Alwis Weerasinghe, du National Central Statistics Office de Namibie. Les effets liés à la conception des enquêtes effectuées au Brésil ont été calculés par M. Pedro Silva, IBGE. Pour les autres enquêtes, les effets liés à la conception ont été calculés par M. Hans Pettersson sur la base des données fournies par les instituts nationaux de statistiques.

Tableau VII.2
Estimation des effets liés à la conception découlant de 7 enquêtes effectuées en Afrique et en Asie du Sud-Est

		Urbain	Rural	National
Enquête sur les dépenses et la consommation en Rép. dém. pop. lao, 1997-1998	Consommation mensuelle totale par ménage	3,8	7,8	5,4
	Consommation alimentaire mensuelle par ménage	4,4	6,8	5,8
	Proportion des ménages ayant accès à un véhicule à moteur	1,3	3,3	2,1
	Proportion des ménages ayant accès à la TV	3,1	6,8	5,4
	Proportion des ménages ayant accès à la radio	2,7	4,8	4,5
	Proportion des ménages ayant accès à la vidéo	3,9	6,1	5,5
Enquête socio-économique au Cambodge, 1999	Consommation mensuelle totale par ménage	2,0	2,0	1,4
	Consommation alimentaire mensuelle par ménage	3,1	3,2	3,2
	Proportion des ménages ayant accès à la TV	2,4	2,2	2,6
Enquête sur le revenu et les dépenses des ménages en Namibie, 1993-1994	Consommation annuelle totale par ménage	2,9	1,9	2,5
	Revenu annuel total par ménage	2,9	2,8	2,8
	Proportion des ménages ayant accès à la TV	6,0	4,6	4,1
	Proportion des ménages ayant accès à la radio	2,7	2,1	2,4
	Proportion des ménages ayant accès au téléphone	6,2	4,6	4,5
Enquête démographique effectuée entre deux recensements en Namibie	Proportion des ménages ayant accès à la TV	14,7	4,1	6,6
	Proportion des ménages utilisant l'électricité pour l'éclairage	4,4	3,9	4,2
	Proportion des ménages ayant eu un décès au cours des 12 derniers mois	2,1	4,3	2,3
Enquête polyvalente sur les ménages au Viet Nam, 1999	Indice de pauvreté	7,1
Enquête sur la main-d'œuvre au Lesotho, 1997	Taux d'emploi	5,6	3,1	6,6
	Proportion de la population âgée de 10 ans et plus qui n'a jamais fréquenté l'école	4,6	5,9	5,5
	Proportion d'agriculteurs pratiquant une agriculture de subsistance	6,3	4,4	8,1
	Proportion de travailleurs indépendants	3,0	1,4	2,4
Enquête d'octobre 1999 sur les ménages en République sud-africaine	Taux d'emploi	4,0	3,6	3,8
Enquête de 2000 sur la main- d'œuvre en République sud-africaine	Taux d'emploi	2,5	3,4	2,8

Note : Deux pointillés (..) indiquent que les données ne sont pas disponibles.

Tableau VII. 3
**Estimations des effets liés à la conception pour certaines estimations
sur les ménages au niveau du pays et par type de zones géographiques (PNAD 1999)**

Variable	National	Zones métropolitaines	Grandes municipalités	Autres zones
Proportion de la population ayant accès à l'approvisionnement en eau	9,80	6,60	6,74	10,73
Proportion de la population ayant accès à l'eau d'une source	9,24	4,04	4,19	9,43
Proportion de la population ayant accès à un assainissement adéquat	9,04	6,36	5,87	11,59
Proportion de la population ayant accès à l'eau de canalisations	8,48	5,16	4,79	9,40
Proportion de la population ayant au moins une salle de bains	8,34	1,51	7,20	7,76
Proportion de la population possédant une terre	8,10	11,53	4,49	7,09
Proportion de la population ayant l'électricité	7,92	1,03	4,43	7,27
Proportion de la population ayant des murs dans un matériau adéquat	7,43	6,17	5,01	6,84
Proportion de la population ayant accès à l'eau de canalisations dans au moins une pièce	7,09	4,74	5,45	7,04
Proportion de la population ayant un toit dans un matériau adéquat	5,68	2,91	2,41	5,65
Nombre moyen de pièces par ménage	5,32	6,26	4,50	5,09
Proportion de la population ayant le téléphone	4,80	5,59	4,44	5,91
Proportion de la population ayant un réfrigérateur	4,59	1,53	2,77	5,02
Proportion de la population ayant une machine à laver	4,34	3,98	3,49	6,25
Proportion de la population ayant un téléviseur couleurs	4,31	1,77	2,76	4,88
Proportion de la population ayant un congélateur	3,83	3,55	2,68	4,67
Proportion de la population ayant un filtre à eau	3,39	2,50	2,07	4,37
Proportion de la population ayant la radio	3,01	1,46	1,62	3,29
Proportion de la population ayant une télévision en noir et blanc	2,79	1,50	1,30	2,93
Loyer moyen	2,52	3,09	2,01	3,39
Proportion de ménages propriétaires de leur logement	2,46	3,18	1,74	2,30
Proportion de ménages vivant dans un logement en location	2,32	2,71	1,78	2,51
Nombre moyen de pièces utilisées comme chambres à coucher	2,14	2,37	1,72	2,09

14. Leurs effets liés à la conception varient entre 2 et 10 pour les estimations au niveau national, avec une moyenne de 5,5. Ces effets sont plus élevés pour les variables telles que la proportion de ménages ayant accès à l'approvisionnement en eau, à l'eau d'une source ou à un assainissement adéquat. Cela était à prévoir, étant donné le degré élevé de mise en grappes que ces variables tendent à afficher. Ils sont plus faibles pour certaines variables « économiques » telles que le loyer moyen, la proportion des ménages possédant leur logement et celle des ménages vivant dans un logement en location, et le nombre moyen de pièces utilisées comme chambres à coucher. Comme il était également à prévoir, les effets liés à la conception sont généralement plus faibles pour les zones métropolitaines et les grandes municipalités, où la conception de l'échantillonnage est en grappes à deux phases, que pour les autres zones, où il est plus agglutiné (échantillonnage en grappes à trois phases).

15. Les effets liés à l'échantillonnage pour un ensemble de variables mesurées au niveau des personnes sont présentés au tableau VII.4.

16. Les effets liés à la conception pour les estimations au niveau national varient d'environ 2 à 16, avec une moyenne de 6,2. Ils sont très élevés pour les variables concernant la race; ils sont élevés pour les variables concernant l'emploi ou le revenu; et ils sont faibles pour les variables telles que la proportion de la population fréquentant l'école et la proportion de la population percevant des prestations d'éducation. Une fois encore, ils sont plus élevés pour les autres zones où la conception est à trois phases. Ils sont généralement plus faibles pour les variables concernant les ménages que pour les variables au niveau des personnes, ce qui n'est pas étonnant vu que le nombre de

personnes est évidemment plus élevé que le nombre de ménages interrogés par UPE. Ces variations substantielles des effets liés à la conception pour différentes variables sont normales car celles-ci affichent des degrés différents de mise en grappes. Ces effets assez élevés s'expliquent aussi pour l'utilisation d'une allocation disproportionnée de l'échantillon entre les strates, qui se traduisent par des pondérations différentes.

Tableau VII.4

Estimations des effets liés à la conception pour certaines caractéristiques mesurées au niveau des personnes à l'échelon national et pour divers sous-domaines (PNAD 1999)

Variable	National	Zones métropolitaines	Grandes municipalités	Autres zones
Proportion de la population de race blanche	15,97	11,97	8,14	19,97
Proportion de la population de race noire ou de couleur	15,75	12,23	8,44	19,41
Proportion de travailleurs rémunérés	8,44	4,45	5,81	7,49
Proportion de travailleurs indépendants	7,65	3,73	5,51	6,66
Proportion de la population affiliée à la sécurité sociale	6,59	2,93	3,28	8,45
Proportion de la population illettrée	6,33	3,67	4,37	7,10
Revenu moyen de la profession principale	5,54	7,16	4,45	6,38
Proportion de la population percevant une allocation logement	5,23	3,80	3,00	5,54
Proportion de la population percevant une prime de transport	4,93	2,94	2,78	9,10
Proportion de la population ayant une assurance maladie	4,90	3,76	2,29	8,79
Proportion de la population active (10 ans ou +)	4,79	1,97	1,67	7,08
Proportion de la population percevant une aide alimentaire	3,35	2,60	2,08	4,60
Proportion d'enfants de 5 à 9 ans qui travaillent	3,27	1,25	2,04	3,00
Proportion d'employeurs	2,87	2,80	1,54	2,63
Proportion de la population scolarisée	1,88	1,75	1,57	1,94
Proportion de la population percevant des prestations d'éducation	1,87	1,85	1,74	2,22

17. Les effets liés à la conception pour l'enquête PME au Brésil sont présentés au tableau VII.5 pour une sélection des estimations publiées chaque mois. Ces valeurs ont été obtenues pour septembre 1999, mois qui a été choisi car la période de référence est la même que pour l'enquête PNAD 1999.

Tableau VII.5

Estimation des effets liés à la conception pour certaines estimations de l'enquête PME pour septembre 1999

Variable	Recife	Salvador	Belo Horizonte	Rio de Janeiro	São Paulo	Pôrto Alegre	Ensemble
Revenu moyen, principale profession	3,43	4,47	2,49	4,44	4,89	4,79	6,23
Proportion employeurs	2,00	2,16	3,06	2,53	2,33	2,27	3,34
Proportion d'illettrés	4,23	4,43	1,86	2,69	2,11	2,13	3,24
Taux de chômage	1,64	2,62	1,98	2,06	1,65	1,67	2,43
Proportion de personnes ayant un emploi déclaré	1,61	1,87	1,66	1,50	1,40	1,75	2,02
Proportion de la population économiquement active	1,59	1,99	1,78	1,61	1,31	1,40	1,96
Proportion de travailleurs rémunérés	1,51	1,67	1,43	1,37	1,34	1,55	1,88
Proportion de travailleurs indépendants	1,53	2,26	1,60	1,47	1,19	1,14	1,78
Proportion de la population scolarisée	1,41	1,57	1,64	1,24	1,26	1,49	1,72

Tableau VII.6
**Estimation des effets liés à la conception pour certaines estimations
 de l'enquête PPV**

Estimations des paramètres de population	Estimation des effets liés à la conception estimée
Nombre de personnes âgées de plus de 14 ans qui sont illettrées	4,17
Proportion des personnes âgées de plus de 14 ans qui sont illettrées	3,86
Nombre de personnes ayant indiqué un « mauvais » état de santé	3,37
Proportion de ménages vivant dans un logement en location	2,97
Nombre moyen de personnes par ménage	2,64
Nombre de personnes âgées de 7 à 14 ans qui sont illettrées	2,64
Proportion des personnes âgées de 7 à 14 ans qui sont illettrées	2,46
Nombre de femmes âgées de 12 à 49 ans qui ont eu des enfants mort-nés	2,03
Nombre de femmes âgées de 12 à 49 ans qui ont eu des enfants	2,02
Nombre de femmes âgées de 12 à 49 ans qui ont eu des enfants nés vivants	2,02
Taux de dépendance (nombre de personnes âgées de 0 à 14 ans plus nombre de personnes âgées de 65 ans et plus, divisé par le nombre de personnes âgées de 15 à 64 ans)	1,99
Nombre moyen d'enfants nés par femme âgée de 12 à 49 ans	1,26

18. Bien que n'étant pas indiqués ici, les effets liés à la conception pour les mêmes estimations ont été calculés pour d'autres mois, et l'on a constaté très peu de variations d'un mois à un autre. L'échantillon de zones d'énumération est fixé pour toute la décennie et la taille des échantillons varie peu sur de courtes périodes. Les effets liés à la conception sont plus importants pour le revenu moyen dans la principale profession et ne sont que modérés pour la proportion de la population illettrée et la proportion d'employeurs. Il n'est pas étonnant que cela soit conforme aux valeurs observées pour des estimations similaires calculées à partir de l'enquête PNAD pour les zones métropolitaines, car on a utilisé essentiellement le même type d'échantillon pour les enquêtes PME et PNAD, à l'exception du fait que l'échantillon prélevé par UPE était plus grand pour l'enquête PME. Pour les autres variables, les effets liés à la conception sont inférieurs à 2,5. Les effets observés pour des variables de l'enquête PME, généralement plus faibles que pour les variables comparables de l'enquête PNAD, s'expliquent par le fait que l'allocation de l'échantillon est plus proche d'une allocation proportionnelle pour la première que pour la seconde.

19. Les effets liés à la conception pour l'enquête PPV au Brésil sont présentés au tableau VII.6 pour une petite sélection des estimations de cette enquête.

20. Pour les estimations considérées ici, les effets liés à la conception varient entre 1,3 et 4,2. Les valeurs relativement faibles de ces effets s'expliquent par le plus faible degré de mise en grappes pour l'enquête PPV, où seuls 8 ménages ont été sélectionnés par UPE. Elles s'expliquent aussi par le fait que la plupart des variables des blocs « démographie » et « éducation » du questionnaire ont été considérées, plus deux variables au niveau des ménages.

21. Nous choisirons à présent, à partir des tableaux VII.2 à VII.6 un ensemble d'estimations qui apparaissent dans plusieurs enquêtes. Les effets liés à la conception sont présentés au tableau VII.7. Ils ont été groupés en trois catégories : a) consommation et revenus des ménages; b) possession de bien durables par les ménages; et c) emploi et profession. Dans chaque catégorie, nous avons groupé les estimations qui ont sensiblement la même définition.

Tableau VII.7
Comparaisons des effets liés à la conception observés
dans les différentes enquêtes

Thème/caractéristique	Urbain	Rural	National	Commentaires
Consommation, revenu des ménages (variables concernant les ménages)				
• Consommation mensuelle totale (République démocratique populaire lao : LECS)	3,8	7,7	5,4	
• Consommation mensuelle totale (Cambodge : CSES)	2,0	2,0	1,4	
• Consommation totale des ménages (Namibie : NHIES)	2,9	1,9	2,5	La taille des grappes de l'enquête CSES est la moitié de celle des grappes des enquêtes LECS et NHIES
• Consommation alimentaire mensuelle (République démocratique populaire lao : LECS)	4,4	6,8	5,8	
• Consommation alimentaire mensuelle (Cambodge : CSES)	2,5	3,3	3,3	
Biens de consommation durables des ménages (variables concernant les ménages)				
• Proportion de ménages ayant accès à la télévision (République démocratique populaire lao : LECS)	3,1	6,8	5,4	
• Proportion de ménages ayant accès à la télévision (Cambodge : CSES)	2,4	2,2	2,6	
• Proportion de ménages ayant accès à la télévision (Namibie : NHIES)	6,0	4,6	4,1	
• Proportion de ménages ayant accès à la télévision (Namibie : NIDS)	14,7	4,1	6,6	Le fait que la taille des grappes de l'enquête NIDS est plus du double de celle des grappes des autres enquêtes explique l'ampleur de l'effet lié à la conception dans les zones urbaines (mais pas le faible niveau de cet effet dans les zones rurales)
• Proportion de ménages possédant un téléviseur couleur (Brésil : PNAD)	4,3	
• Proportion de ménages ayant accès à la radio (République démocratique populaire lao : LECS)	2,7	4,8	4,5	
• Proportion de ménages ayant accès à la radio (Cambodge : CSES)	2,1	2,8	3,4	
• Proportion de ménages ayant accès à la radio (Namibie : NHIES)	2,7	2,1	2,4	
• Proportion de ménages ayant accès au téléphone (Namibie : NHIES)	6,2	4,6	4,5	
• Proportion de ménages ayant accès au téléphone (Brésil : PNAD)	—	—	4,8	
Emploi, profession (variables concernant les personnes)				
• Taux d'emploi (Afrique du Sud : OHS)	4,0	3,6	3,8	La différence de niveau d'effets liés à la conception pour les zones urbaines entre l'Afrique du Sud (LFS) et l'Afrique du Sud (OHS) est due à la plus petite taille des grappes dans le domaine urbain pour l'enquête LFS (5 ménages contre 10 pour l'enquête OHS)
• Taux d'emploi (Afrique du Sud : LFS)	2,5	3,4	2,8	
• Taux d'emploi (Lesotho : LFS)	5,6	3,1	6,6	
• Taux d'emploi (Brésil : PNAD)	—	—	4,8	

Note : Les deux pointillés (..) indiquent que les données ne sont pas connues.
Le tiret (—) indique que ce point est sans objet.

22. Les effets liés à la conception pour les estimations à l'échelon national varient entre 1,4 et 6,6 avec une valeur médiane de 4,3. Certains de ces effets sont très élevés. Celui qui se détache est l'effet de 14,7 pour la proportion de ménages urbains ayant accès à la télévision dans l'enquête NIDS en Namibie. Ce niveau élevé s'explique par la taille des grappes de 50 ménages : si cette taille avait été de 20 comme dans

l'enquête NHIES, l'effet lié à la conception aurait été de 6,7, soit proche de celui de 6 de cette enquête. Ce niveau reste néanmoins élevé, et les pondérations des variables n'apportent pas une contribution appréciable à cet égard. Les effets liés à la conception pour la plupart des estimations rurales de l'enquête LECS sont eux aussi élevés. Dans l'enquête NHIES, les effets liés à la conception sont également élevés dans certaines estimations urbaines pour les biens de consommation durables.

23. Dans toutes les enquêtes sauf les deux enquêtes effectuées en Afrique du Sud et l'enquête effectuée au Cambodge, on note de nettes différences entre zones urbaines et rurales. Dans les enquêtes effectuées en République démocratique populaire lao et au Brésil (voir tableaux VII.2 à VII.6), les effets liés à la conception sont généralement plus faibles en milieu urbain qu'en milieu rural. Dans les enquêtes effectuées en Namibie et au Lesotho, c'est l'inverse. (Pour la plupart des enquêtes, la taille des grappes était la même dans les zones urbaines et rurales, de sorte que les différences notées ne sont pas dues à des différences de taille des grappes.)

24. Les effets liés à la conception comprennent les effets de la stratification, de pondérations inégales et de la taille et de l'homogénéité des grappes (voir l'analyse détaillée de ces effets au chapitre VI). Les enquêtes présentées au tableau VII.7 peuvent être similaires dans l'ensemble dans la conception de leurs échantillons, mais elles présentent de nettes différences de stratification, de taille des grappes, d'allocation de l'échantillon, etc. Il est donc difficile de comparer les effets liés à la conception entre ces enquêtes, y compris pour la même estimation. Pour les rendre plus comparables, il faut retirer des effets liés à la conception ceux qui résultent de la taille des grappes et des pondérations.

D. Calcul des taux d'homogénéité

25. On peut poursuivre cette analyse sur un plus petit ensemble d'enquêtes et de variables, en utilisant quelques estimations de la consommation des ménages et de la possession de biens de consommation durables des enquêtes LECS, CSES et NHIES, trois enquêtes qui ont des échantillons de conception similaire. Toutes ces enquêtes ont utilisé des modèles d'échantillon à deux phases avec les ZE comme unités primaires d'échantillonnage. Les UPE ont été stratifiées sensiblement de la même manière entre les provinces et entre les divisions urbaines/rurales à l'intérieur des provinces. Les ménages ont été sélectionnés par échantillonnage systématique à l'intérieur des ZE. En revanche, l'allocation de l'échantillon entre les strates était différente. Entre les provinces, l'enquête lao avait une allocation égale, alors que les deux autres enquêtes avaient des allocations presque proportionnelles. Cette analyse a pour objet d'examiner l'effet d'échantillons de conception complexe sur la précision d'estimations (approximativement) semblables sur différentes populations et d'étudier les similarités et d'éventuelles tendances des taux d'homogénéité.

26. Dans un premier temps, on retire les effets de pondérations inégales des effets liés à la conception. Au tableau VII.8, on a séparé les effets liés à la conception en composantes pondération et mise en grappes. On calcule ces composantes en utilisant les équations 23 et 30 du chapitre VI. Les échantillons de tailles égales à l'intérieur des provinces dans l'enquête LECS produisent une variation sensible des pondérations de l'échantillonnage. De ce fait, les effets liés à la conception résultant de la pondération sont assez élevés pour les estimations de cette enquête. L'enquête NHIES présente un certain degré de suréchantillonnage dans les régions moins peuplées et dans les zones urbaines, ce qui explique pourquoi les effets liés à la conception résultant de la pondération sont supérieurs à 1,0 mais considérablement plus faibles que les effets observés

dans l'enquête LECS. L'enquête CSES se caractérise par un suréchantillonnage dans les zones urbaines.

Tableau VII.8

Ensemble des effets liés à la conception divisés en effets résultant de la pondération $d_w^2(\bar{y})$ et de la mise en grappes $d_{cl}^2(\bar{y})$,

Statistique/caractéristique	Urbain			Rural		
	Ensemble $d^2(\bar{y})$	Pondération $d_w^2(\bar{y})$	Mise en grappes $d_{cl}^2(\bar{y})$	Ensemble $d^2(\bar{y})$	Pondération $d_w^2(\bar{y})$	Mise en grappes $d_{cl}^2(\bar{y})$
Consommation, revenu des ménages						
• Consommation mensuelle totale (LECS)	3,8	1,60	2,4	7,7	1,55	5,0
• Consommation mensuelle totale (CSES)	2,0	1,11	1,8	2,0	1,16	1,7
• Consommation totale des ménages (NHIES)	2,9	1,20	2,4	1,9	1,23	1,5
• Consommation alimentaire mensuelle (LECS)	4,4	1,60	2,8	6,8	1,55	4,4
• Consommation alimentaire mensuelle (CSES)	2,5	1,11	2,3	3,3	1,16	2,8
• Revenu total des ménages (NHIES)	2,9	1,20	2,4	2,8	1,23	2,3
Biens de consommation durables des ménages						
• Proportion de ménages ayant accès à la télévision (LECS)	3,1	1,60	2,0	6,8	1,55	4,4
• Proportion de ménages ayant accès à la télévision (CSES)	1,9	1,11	1,7	1,8	1,16	1,6
• Proportion de ménages ayant accès à la télévision (NHIES)	6,0	1,20	5,0	4,6	1,23	3,7
• Proportion de ménages ayant accès à la radio (LECS)	2,7	1,60	1,7	4,8	1,55	3,1
• Proportion de ménages ayant accès à la radio (CSES)	2,1	1,11	1,9	2,3	1,16	2,0
• Proportion de ménages ayant accès à la radio (NHIES)	2,7	1,20	2,3	2,1	1,23	1,7
• Proportion de ménages ayant accès à la vidéo (LECS)	3,9	1,60	2,4	6,1	1,55	3,9
• Proportion de ménages ayant accès au téléphone (NHIES)	6,2	1,20	5,2	4,6	1,23	3,7

27. Ces trois enquêtes ont utilisé une conception selon laquelle un nombre constant de ménages a été sélectionné dans chaque UPE (échantillonnage systématique). La taille constante des grappes contribue également à la variation des pondérations car des imperfections dans la mesure de la taille des UPE entraînent une variation de l'ensemble des pondérations de l'échantillonnage.

28. Les effets liés à la conception résultant de la mise en grappes, $d_{cl}^2(\bar{y})$, dépendent de la taille de l'échantillon des grappes. Les enquêtes effectuées en République démocratique populaire lao et en Namibie avaient des échantillons de 20 ménages par grappe, tandis que l'enquête effectuée au Cambodge avait un échantillon de 10 ménages par grappe. Pour supprimer les effets résultant des différentes tailles des grappes lorsque l'on compare les résultats des différentes enquêtes, nous avons calculé des taux d'homogénéité (*roh*) pour les estimations du tableau VII.8 (voir équation 30 au chapitre VI). Les résultats de ces calculs sont présentés au tableau VII.9. Les *roh* mesurent l'homogénéité interne des UPE (zones d'énumération) pour les variables de l'enquête. La question à examiner est de savoir s'il existe des similarités dans les niveaux et les profils du *roh* à travers les pays.

Tableau VII.9
Taux d'homogénéité pour les domaines urbain et rural

Statistique/caractéristique	Urbain	Rural	Ratio urbain/rural
Consommation, revenu des ménages			
• Consommation mensuelle totale (LECS)	0,072	0,209	0,3
• Consommation mensuelle totale (CSES)	0,089	0,080	1,1
• Consommation totale des ménages (NHIES)	0,071	0,025	2,9
• Consommation alimentaire mensuelle (LECS)	0,092	0,178	0,5
• Consommation alimentaire mensuelle (CSES)	0,139	0,204	0,7
• Revenu total des ménages (NHIES)	0,071	0,058	1,2
Biens de consommation durables des ménages			
• Accès à la télévision (LECS)	0,049	0,178	0,3
• Accès à la télévision (CSES)	0,079	0,061	1,3
• Accès à la télévision (NHIES)	0,200	0,125	1,6
• Accès à la radio (LECS)	0,036	0,110	0,3
• Accès à la radio (CSES)	0,100	0,109	0,9
• Accès à la radio (NHIES)	0,063	0,032	1,9
• Proportion de ménages ayant accès à la vidéo (LECS)	0,076	0,154	0,5
• Accès au téléphone (NHIES)	0,208	0,125	1,7

29. Comme l'homogénéité des grappes peut être différente entre les grappes urbaines et rurales, la valeur du *roh* a été calculée séparément pour ces deux catégories de population. Les résultats sont présentés au tableau VII.9. Certains résultats de ce tableau se distinguent des autres.

- Les profils des différences entre les valeurs du *roh* pour les zones urbaines et rurales sont différents dans les trois pays. Les valeurs du *roh* pour les grappes urbaines de l'enquête effectuée en République démocratique populaire lao sont régulièrement plus faibles que les valeurs correspondantes pour les grappes rurales. Le ratio moyen urbain/rural est de 0,4. Dans l'enquête effectuée en Namibie, les différences vont dans le sens inverse : les valeurs du *roh* pour les grappes urbaines sont en moyenne plus élevées que les valeurs correspondantes pour les grappes rurales, selon un facteur de 1,9. Dans l'enquête effectuée au Cambodge, on ne distingue pas clairement de tendance pour le *roh* entre les zones urbaines et rurales.
- Les *roh* correspondant aux grappes rurales sont élevés dans l'enquête LECS (ils se situent dans une fourchette de 0,110 à 0,209, avec une valeur médiane de 0,178). Les *roh* correspondant aux grappes urbaines sont nettement plus faibles (ils se situent dans une fourchette de 0,036 à 0,092, avec une valeur médiane de 0,072).
- Le *roh* correspondant à la consommation alimentaire mondiale est élevé dans les zones rurales au Cambodge (0,204). Ce *roh* est considérablement plus élevé pour la consommation mensuelle totale et également plus élevé pour les estimations de biens de consommation durables des ménages.

30. Les fortes différences observées entre les *roh* des zones rurales et urbaines en République démocratique populaire lao sont dues principalement aux valeurs

élevés du *roh* dans les zones rurales. Ces résultats sont conformes à ceux d'une précédente enquête du même type effectuée dans le pays. Les valeurs élevées du *roh* enregistrées dans les zones rurales ne sont pas déraisonnables, étant donné que les villages ruraux sont petits et assez homogènes du point de vue socio-économique. Le caractère saisonnier que présentent les chiffres de la consommation mensuelle totale et de la consommation alimentaire mensuelle sont peut-être autant de facteurs qui expliquent ces variables. Chaque UPE est visitée pendant un mois et les visites à l'échantillon d'UPE sont étalées sur une période de 12 mois. Par conséquent, il se produit une « mise en grappes saisonnière » en plus de la mise en grappes géographique. Il y a des raisons de croire que ce caractère saisonnier est sensiblement plus prononcé dans les zones rurales.

31. En Namibie, un grand nombre des UPE rurales sont situées dans des zones d'agriculture de rapport et sont donc assez hétérogènes, en ce sens qu'elles englobent un ensemble de familles d'agriculteurs à revenu élevé et de familles d'ouvriers agricoles à faible revenu. Dans les zones urbaines, en revanche, on note une assez forte différenciation fondée sur le niveau de revenu dont il n'a été tenu compte que partiellement dans la stratification. Ces conditions sont peut-être ce qui explique que les *roh* sont plus élevés pour la consommation et le revenu des ménages dans les zones urbaines.

32. Il convient d'ajouter deux autres explications à celles données ci-dessus. La première, c'est que les effets liés à la conception (et, par conséquent, le *roh*) pour les variables concernant la consommation sont assez sensibles aux valeurs à l'extrémité supérieure. Dans certains cas, en retirant quelques-unes des valeurs les plus élevées, on change considérablement l'effet lié à la conception. La seconde, c'est que les valeurs du *roh* reflètent plus que simplement les mesures de l'homogénéité des grappes. Elles saisissent aussi les effets de la variance liée à l'enquêteur, lorsque divers enquêteurs, ou équipes d'enquêteurs, effectuent les entrevues dans différentes UPE.

E. Discussion

33. Le tableau VII.9 ne permet pas de discerner de similarités entre les pays en ce qui concerne les niveaux ou profils du *roh*. Ces résultats offrent une maigre consolation pour le statisticien par échantillonnage qui cherche à utiliser le *roh* d'une enquête similaire effectuée dans un autre pays lorsqu'il conçoit l'échantillon à utiliser pour son enquête. Il semble que les conditions propres à la population du pays contribuent pour beaucoup au degré d'homogénéité des grappes pour les types de variables socio-économiques étudiées ici. Certes, cette étude est très limitée; la seule conclusion générale que l'on peut en tirer est qu'il faut faire preuve de prudence lorsque l'on « importe » le *roh* d'une enquête effectuée dans un autre pays. Les résultats font également ressortir la nécessité de calculer et de documenter les effets liés à la conception et les *roh* d'une enquête afin qu'ils puissent servir lors de la conception de la suivante.

34. Néanmoins, les résultats de cette étude, pour incertains qu'ils soient, vont à l'encontre des conclusions habituelles. Les études des Enquêtes démographiques et sanitaires ont révélé que les estimations du *roh* d'une enquête donnée sont relativement transposables à travers les pays, à conditions que les échantillons soient de conception comparable (voir chapitre XII). De même, l'étude effectuée sur un certain nombre d'enquêtes mondiales sur la fécondité a conclu à des similarités dans les profils du *roh* d'un pays à un autre. Il se peut que le *roh* pour les variables démographiques « se comporte mieux » et soit plus facilement transposable que les *roh* des variables socio-économiques.

Annexe

Description des échantillons conçus pour les 11 enquêtes sur les ménages

Les échantillons conçus pour ces 11 enquêtes sont décrits brièvement ci-dessous :

Enquête de 1997-1998 sur les dépenses et la consommation en République démocratique populaire lao (LECS)

Les zones d'énumération (ZE) du recensement ont été utilisées comme UPE. Les UPE ont été stratifiées en 18 provinces et en zones urbaines et rurales. Les ZE rurales ont été à nouveau stratifiées selon qu'elles avaient ou non « une route d'accès ». Des échantillons de taille égale de 25 UPE ont été sélectionnés systématiquement avec probabilité proportionnelle à la taille [*Probability proportional to size* (PPS)] dans chaque province (en tout, 450 UPE) [Rosen, 1997]. Vingt ménages ont été sélectionnés dans chaque UPE, produisant un échantillon de 9 000 ménages. L'allocation de l'échantillon à égalité entre toutes les provinces s'est traduite par une forte variation des pondérations d'échantillonnage au niveau des ménages.

Enquête socio-économique au Cambodge 1999 (CSES)

Les villages ont servi d'UPE. Quelques communes et villages ont été exclus parce qu'ils ne pouvaient recevoir la visite d'enquêteurs pour des raisons de sécurité; ces zones exclues comptaient 3,4 % du nombre total des ménages du pays.

Les villages ont été groupés en cinq strates sur la base de zones écologiques. Phnom Penh a été traitée comme une strate distincte, et les secteurs ruraux et urbains ont été également traités comme des strates distinctes. Ainsi, 10 strates ont été créées à partir des quatre zones géographiques (Phnom Penh, les Plaines, Tonle Sap, Zone côtière et Plateau/Montagne). À partir de chaque strate, on a extrait quatre sous-échantillons indépendants de villages. L'échantillon a été alloué de façon approximativement proportionnelle aux différentes strates.

Six cents villages ont été sélectionnés par échantillonnage PPS systématique circulaire. Dix ménages ont été sélectionnés à l'intérieur de chaque village (Institut national de statistiques, Royaume du Cambodge, 1999).

Enquête sur le revenu et les dépenses des ménages en Namibie 1993-1994 (NHIES)

Les UPE étaient essentiellement les zones d'énumération du recensement. Quelques petites ZE ont été combinées avec des ZE voisines avant la sélection. La taille moyenne des UPE était d'environ 150 ménages. Une stratification primaire a été effectuée selon les divisions urbaine/rurale et 14 régions. Une stratification secondaire a été effectuée dans le domaine urbain où ont été définies des strates « urbaines » et « semi-urbaines ». L'échantillon a été alloué aux différentes strates de façon approximativement proportionnelle. Toutefois, les zones urbaines ont fait l'objet d'un léger suréchantillonnage. On a sélectionné un échantillon de 96 UPE urbaines et de 123 UPE rurales en utilisant une procédure PPS systématique.

Enquête démographique effectuée entre deux recensements en Namibie 1995-1996 (NIDS)

La conception a été la même que celle de l'enquête NHIES. Un échantillon de 82 UPE urbaines et de 120 UPE rurales a été sélectionné. Pour cette enquête, un assez gros échan-

tillon de 50 ménages a été sélectionné dans chaque UPE, donnant un échantillon total de 9 500 ménages (Pettersson, 1997).

Enquête polyvalente sur les ménages au Viet Nam 1999 (VMPHS)

Les communes ont été utilisées comme UPE dans les zones rurales. Dans les zones urbaines, ce sont les quartiers qui ont servi d'UPE. La stratification a été effectuée entre zones urbaines et rurales et entre provinces (61 provinces). En tout, 839 communes ont été sélectionnées par PPS. L'échantillon était essentiellement de la même taille pour chaque province, mais des échantillons un peu plus grands ont été alloués aux grandes provinces. Les unités secondaires d'échantillonnage (USE) étaient les villages à l'intérieur des communes et les blocs à l'intérieur des quartiers. Deux USE ont été sélectionnées à l'intérieur de chaque commune sélectionnée. Dans chaque USE, 15 ménages ont été sélectionnés. En tout, environ 25 000 ménages ont été sélectionnés (Phong, 2001).

Enquête sur la main-d'œuvre au Lesotho 1997

L'échantillon était un échantillon à deux phases. Les unités primaires d'échantillonnage étaient des groupes de zones d'énumération. La taille moyenne des UPE était de 370 ménages. Les UPE ont été stratifiées en divisions urbaines/rurales, régions (10) et zones agroéconomiques, ce qui a produit un total de 33 strates. L'échantillon a été alloué de façon proportionnelle aux strates, à deux exceptions près : deux petites strates ont été fortement suréchantillonnées. Une procédure PPS systématique a été employée pour sélectionner 120 UPE. À l'intérieur des UPE, 15 à 40 ménages ont été sélectionnés, par échantillonnage aléatoire systématique, ce qui a produit un échantillon total de 3 600 ménages. Tous les membres admissibles des ménages ont été inclus dans l'enquête (Pettersson, 2001).

Enquête d'octobre 1999 sur les ménages en République sud-africaine (OHS)

Les zones d'énumération (ZE) du recensement ont été utilisées comme UPE. Pendant le processus de sélection, les ZE qui comptaient moins de 80 ménages ont été groupées avec des ZE voisines sur la liste selon une méthode proposée par Kish (1965). La taille moyenne des UPE était de 80 à 100 ménages pour les UPE urbaines et de 110 à 120 ménages pour les UPE rurales. Les UPE ont été stratifiées en neuf provinces. L'échantillon a été alloué aux strates proportionnellement à leur racine carrée. À l'intérieur de chaque province, une nouvelle stratification a été effectuée selon les conseils de district et les conseils métropolitains. Un échantillon de 2 984 UPE a été sélectionné par échantillonnage PPS systématique : 1 711 dans les zones urbaines et 1 273 dans les zones rurales. Dans chaque UPE, un échantillon systématique de 10 « points de visite » (approximativement le même nombre que celui des ménages) a été prélevé (Stoker, 2001).

Enquête de février 2000 sur la main-d'œuvre en République sud-africaine

L'Enquête de février 2000 sur la main-d'œuvre a été la première enquête à utiliser un nouvel échantillon-maître qui avait été établi à la fin de 1999 à partir de la base de données du recensement de 1996. Cet échantillon était composé de 2 000 UPE (par la suite, la même année, il a été porté à 3 000 UPE). Les zones d'énumération du recensement ont été utilisées comme UPE, et les ZE de moins de 100 ménages ont été groupées avec des ZE voisines. Les UPE ont été stratifiées en neuf provinces. L'échantillon a été alloué aux strates par allocation selon la racine carrée. Dans chaque UPE, des grappes de 10 points de visite ont été formées, chaque grappe étant répartie sur toute l'UPE. Un ensemble de grappes ont été sélectionnées pour la prochaine Enquête sur la main-d'œuvre.

En raison de difficultés budgétaires, il a été décidé de ramener l'enquête sur la main-d'œuvre à 10 000 points de visite. Pour ce faire, on a procédé de la façon suivante : dans toutes les UPE urbaines, on n'a retenu que cinq points de visite à partir de la grappe identifiée. Pour chaque échantillon rural, un sous-échantillon PPS systématique regroupant 50 % des UPE rurales a été prélevé sur l'ensemble d'UPE rurales et, parmi les UPE sélectionnées, la grappe entière identifiée de 10 points de visite a été incorporée à l'échantillon (Stoker, 2001).

PNAD (*Pesquisa Nacional por Amostra de Domicílios*) 1999, Brésil

L'enquête PNAD couvre annuellement un échantillon d'environ 115 000 ménages, représentant l'ensemble du Brésil, sauf les zones rurales du nord (Amazonie). La stratification s'est faite sur une base géographique en 36 strates explicites. Ces 36 strates ont été divisées à raison d'une strate pour chacun des 18 des États, et les neuf strates restantes ont été divisées en deux strates chacune. Une strate a ensuite été formée des UPE situées dans la zone métropolitaine entourant la capitale de chaque État, et une strate a regroupé les autres UPE de l'État. Dans les strates formées par les zones métropolitaines, la conception a été de type échantillonnage en grappes en deux phases, où les UPE étaient les zones d'énumération du recensement, sélectionnées par échantillonnage PPS systématique, dont la taille était égale au nombre de ménages dénombrés lors du dernier recensement. Avant leur sélection, les UPE ont été classées selon un code géographique, ce qui a permis une stratification implicite par municipalité et selon une classification urbaine/rurale.

Dans les strates qui ne correspondaient pas aux zones métropolitaines, les UPE étaient des municipalités. Elles ont été stratifiées par taille et par zone géographique, formant des strates de population approximativement égale (au moyen des données du dernier recensement de population disponible). Deux municipalités (des UPE de ces strates) ont ensuite été sélectionnées dans chaque strate par échantillonnage PPS systématique, leur taille étant mesurée par leur population totale. Avant cette sélection systématique, certaines municipalités ont été déclarées UPE « sûres » en raison de leur nombreuse population, et ont donc été incluses dans l'échantillon de municipalités « sûres ». Dans chaque municipalité sélectionnée, les ZE ont été sélectionnées par échantillonnage PPS systématique, et leur taille était égale au nombre de ménages dénombrés lors du dernier recensement de la population. Au dernier stade de la sélection, les ménages ont été sélectionnés au sein de chaque ZE par échantillonnage systématique à partir de listes mises à jour annuellement. Chaque membre des ménages sélectionnés a été inclus dans l'enquête. Un échantillon cible de 13 ménages aurait dû être sélectionné dans chaque ZE. Toutefois, pour réduire les variations de pondération dues à des mesures de taille périmées, on a utilisé des fractions d'échantillonnage constantes dans chaque ZE au lieu de *tailles* constantes, ce qui a eu pour effet de produire des grappes de tailles différentes.

L'allocation de l'échantillon entre les strates était disproportionnée, et le ratio de la plus forte pondération à la plus faible était approximativement de 8.

PME (*Pesquisa Mensal de Emprego*) pour septembre 1999, Brésil

L'enquête PME est une enquête sur la main-d'œuvre qui couvre un échantillon mensuel d'environ 40 000 ménages vivant dans les six plus grandes zones métropolitaines du Brésil, à partir duquel on détermine les principaux indicateurs actuels de la main-d'œuvre. La conception de cet échantillon est la même que pour l'enquête PNAD dans les strates des zones métropolitaines, sauf en ce qui concerne la grappe cible, qui est de 20 pour l'enquête PME contre 13 pour l'enquête PNAD.

PPV (*Pesquisa de Padrões de Vida*) 1996-1997, Brésil

L'enquête PPV portait sur la mesure des niveaux de vie et a utilisé l'approche élaborée pour les Enquêtes sur la mesure du niveau de vie [*Living Standards Measurement Study*].

(LSMS)] effectuées dans divers pays sous le patronage de la Banque mondiale (Grosh et Muñoz, 1996). L'enquête effectuée au Brésil en 1996-1997 portait sur l'étude d'un grand nombre de caractéristiques démographiques, sociales et économiques à partir d'un échantillon de 4 944 ménages sélectionnés à partir de 554 ZE des régions Nord-Est et Sud-Est du Brésil. L'échantillon était un échantillon en grappes stratifié à deux phases. La stratification s'est faite en deux temps. Premièrement, 10 strates géographiques ont été formées pour identifier les six zones métropolitaines de Fortaleza, Recife, Salvador, Belo Horizonte, Rio de Janeiro et São Paulo, plus quatre autres strates qui couvraient le reste des régions Nord-Est et Sud-Est, subdivisées en zones d'énumération urbaines et rurales. À l'intérieur de chacune de ces 10 strates géographiques, les ZE ont encore été subdivisées en trois strates selon le revenu moyen du chef de ménage enregistré lors du recensement de la population de 1991. Au total, 30 strates ont ainsi été formées.

La taille totale de l'échantillon a été fixée à 554 ZE, soit 278 pour la région Nord-Est et 276 pour la région du Sud-Est. L'allocation des ZE a été effectuée par échantillonnage PPS avec une procédure de remplacement, le nombre de ménages par ZE étant utilisé comme mesure de la taille. Dans chaque ZE urbaine sélectionnée, un échantillon fixe de huit ménages a été sélectionné par échantillonnage aléatoire simple sans remplacement. Dans les ZE rurales, on a pris des échantillons de 16 ménages pour des raisons de rapport coût-efficacité.

Malgré la petite taille de son échantillon, comparé à celui des enquêtes PNAD et PME, l'enquête PPV fournit des informations utiles sur les effets liés à la conception car elle utilise une stratification directe des ZE selon le revenu, ainsi qu'un plus petit nombre de ménages par ZE que les autres enquêtes. Son autre trait distinctif vient du fait que l'estimation n'a utilisé que des pondérations de probabilité de sélection inverses et qu'aucun étalonnage n'a été tenté sur les projections de population. Les différences de pondération des échantillons de l'enquête PPV étaient substantielles, la pondération la plus forte étant plus de 40 fois supérieure à la plus faible.

RÉFÉRENCES

- Grosh, M., et J. Muñoz (1996). *A Manual for Planning and Implementing the Living Standards Measurement Study Survey*. Living Standards Measurement Study, document de travail n° 126. Washington, D.C., Banque mondiale.
- Kish, L. (1965). *Survey Sampling*. New York, Wiley.
- National Institute of Statistics, Kingdom of Cambodia (1999). *Cambodia Socio-Economic Survey 1999: Technical Report on Survey Design and Implementation*. Phnom Penh.
- Pettersson, H. (1994). *Master Sample Design: Report from a Mission to the National Central Statistics Office, Namibia, May 1994*. International Consulting Office, Statistics Sweden.
- _____ (1997). *Evaluation of the Performance of the Master Sample 1992-96: Report from a Mission to the National Central Statistics Office, Namibia, May 1997*. International Consulting Office, Statistics Sweden.
- _____ (2001). *Sample Design for Household and Business Surveys: Report from a Mission to the Bureau of Statistics, Lesotho May 21-June 2, 2001*. International Consulting Office, Statistics Sweden.
- Phong, N. (2001). Personal correspondence concerning sample design for the Viet Nam Multipurpose Household Survey 1999.
- Rosen, B. (1997). *Creation of the 1997 Lao Master Sample. Report from a Mission to the National Statistics Centre, Lao PDR*. International Consulting Office, Statistics Sweden.

- Stoker, D. (2001). Personal correspondence concerning sample design for the October Household Survey and Labour Force Survey in the Republic of South Africa.
- Verma, V., C. Scott et C. O'Muirheartaigh (1980). Sample designs and Sampling Errors for the World Fertility Survey. *Journal of the Royal Statistical Society, Série A*, vol. 143, partie 4, pp. 431-473.

Section C

**Erreurs autres
que les erreurs
d'échantillonnage**

Introduction

JAMES LEPKOWSKI

Université du Michigan
Ann Arbor, Michigan
États-Unis d'Amérique

1. Dans les sections et chapitres précédents de la présente publication, nous avons examiné principalement les erreurs d'échantillonnage qui se produisent lorsqu'un échantillon à probabilité représentative est prélevé sur une population. Dans la présente section, nous examinerons un certain nombre d'autres erreurs qui se produisent dans les enquêtes sur les ménages. Comme les erreurs d'échantillonnage, certaines de ces erreurs varient selon les échantillons ou lors de répétitions éventuelles du processus de mesure. D'autres sont fixes, ou systématiques, et ne varient pas d'un échantillon au suivant.

2. Dans le cadre de conception de l'échantillon, les erreurs variables sont généralement appelées variance d'échantillonnage. Il y a les erreurs d'échantillonnage fixes, dont certaines ont déjà été mentionnées, qui sont appelées distorsions. Par exemple, l'exclusion délibérée d'un sous-groupe de population entraîne la non-couverture de ce sous-groupe, c'est-à-dire une erreur qui sera présente, et de la même taille, quel que soit l'échantillon éventuel sélectionné.

3. Les erreurs autres que les erreurs d'échantillonnage comprennent les erreurs de non-observation lorsque l'on n'obtient pas les données d'une unité d'échantillonnage ou d'une variable, ou les erreurs de mesure qui se produisent lorsque l'on collecte les valeurs des variables de l'enquête. Les erreurs de non-observation sont généralement fixes par nature, et incitent à penser qu'il y a des distorsions dans les estimations de l'enquête. Les erreurs de mesure sont parfois fixes, mais elles peuvent aussi être variables.

4. Parmi les erreurs de non-observation, deux sources d'erreurs sont les plus importantes : la non-couverture et la non-réponse. Dans l'échantillonnage par probabilité, il faut qu'il y ait une population composée d'éléments bien définis, dont chacun a une chance de sélection différente de zéro. La non-couverture se produit lorsqu'un élément de la population n'a effectivement aucune chance d'être sélectionné, aucun moyen de faire partie de l'échantillon sélectionné. Cela peut se produire parce qu'un ménage ou une personne refuse catégoriquement de coopérer, ou en raison d'une barrière de langage, d'une incapacité pour raison de santé ou du fait que personne n'est à la maison pendant la période de l'enquête.

5. Les erreurs de mesure peuvent avoir des sources diverses — les déclarants, les enquêteurs, les superviseurs, voire les systèmes de traitement de l'information. Les erreurs de mesure imputables aux déclarants peuvent se produire lorsqu'un déclarant oublie l'information nécessaire et donne une réponse incorrecte, ou fausse les informations qu'il donne en réponse à une question délicate. Les erreurs imputables aux déclarants risquent d'entraîner une distorsion, car le déclarant oublie régulièrement

ou déforme une réponse de la même manière, à quelque moment qu'on lui pose la question. Ces erreurs peuvent aussi être variables. Certains déclarants peuvent oublier une réponse à un moment et s'en souvenir à un autre.

6. Les concepteurs de l'enquête prennent en considération quatre dimensions lorsqu'ils examinent ces types d'erreurs. La première de ces dimensions a trait à la nécessité de bien définir l'erreur et d'examiner ses sources dans le processus de l'enquête, notamment de déterminer quelle partie de ce processus semble engendrer ce type d'erreur. La seconde a trait à la question de savoir comment mesurer l'ampleur de l'erreur, ce qui est particulièrement difficile. La troisième concerne les procédures à élaborer pour réduire l'ampleur de l'erreur, quoique l'application de ces procédures nécessite souvent des ressources additionnelles. Enfin, chaque enquête s'accompagne d'erreurs autres que des erreurs d'échantillonnage, et les concepteurs s'efforcent de les compenser dans les résultats de l'enquête.

7. Les chapitres VIII et IX de la présente section sont consacrés respectivement à l'examen, d'un point de vue théorique, des erreurs de non-observation et de mesure, et offrent des illustrations de plusieurs types différents de ces erreurs. Les chapitres X et XI traitent ces erreurs de façon plus détaillée : le premier examine leur impact général sur la qualité des résultats de l'enquête, et le second présente une étude de cas des ces types d'erreurs dans un pays, le Brésil.

Chapitre VIII

L'erreur de non-observation dans les enquêtes sur les ménages dans les pays en développement

JAMES LEPKOWSKI

Université du Michigan
Ann Arbor, Michigan, États-Unis d'Amérique

RÉSUMÉ

Dans une enquête, il y a non-observation lorsque des mesures ne sont pas ou ne peuvent pas être effectuées sur une partie de la population visée ou de l'échantillon. La non-observation peut être complète, auquel cas aucune mesure n'est effectuée sur une unité (par exemple, un ménage ou une personne), ou partielle, auquel cas certaines mesures désirées, mais pas toutes, sont effectuées sur une unité. Dans le présent chapitre, nous examinerons deux sources de non-observation, la non-couverture et la non-réponse. Il y a non-couverture lorsque des unités étudiées n'ont aucune chance d'être sélectionnées pour l'enquête. Il y a non-réponse lorsqu'une famille ou une personne sélectionnée ne participe pas à celle-ci ou y participe mais fournit des informations incomplètes. Dans ce chapitre, nous examinerons les causes et les conséquences des erreurs de non-observation et les mesures à prendre pour y remédier. La non-couverture et la non-réponse peuvent fausser les estimations de l'enquête lorsque la partie de la population ou de l'échantillon laissée en dehors est différente de la partie observée. Comme ces distorsions peuvent être graves, nous analyserons un certain nombre de remèdes et d'ajustements envisagés.

Termes clés : non-réponse, non-couverture, distorsion, population visée, cadre d'échantillonnage, taux de réponse.

A. Introduction

1. Dans une enquête, la non-observation est le résultat de l'incapacité à effectuer des mesures sur une partie de la population visée. Cette incapacité peut être complète, auquel cas aucune mesure n'est effectuée, ou partielle, auquel cas certaines des mesures désirées sont effectuées, mais pas toutes.

2. L'une des sources évidentes de la non-observation est le processus d'échantillonnage. Le seul cas où il n'y a pas de non-observation découlant du prélèvement d'un échantillon est celui du recensement, où les mesures sont effectuées sur chaque élément de la population. La non-observation découlant de l'échantillonnage donne lieu aux erreurs d'échantillonnage examinées aux chapitres VI et VII de la présente publication. Par conséquent, cette source de non-observation ne sera pas traitée ici.

3. Dans le présent chapitre, nous examinerons deux autres sources de non-observation, à savoir la non-couverture et la non-réponse. Comme nous le verrons de façon plus détaillée par la suite, il y a non-couverture lorsque des unités de la population étudiée n'ont aucune chance d'être sélectionnées pour l'enquête, et il y a non-réponse lorsqu'une unité sélectionnée ne participe pas à l'enquête, que ce soit partiellement ou totalement. Ce chapitre portera sur les causes de ces sources de non-observation, leurs conséquences potentielles, les mesures à prendre pour les limiter et les méthodes employées pour tenter d'atténuer la distorsion qu'elles peuvent engendrer dans les estimations de l'enquête. Parmi les conséquences de la non-couverture figure la possibilité de distorsions dans les résultats de l'enquête. Si une partie de la population qui est laissée en dehors est différente de la partie observée, il y aura des différences entre les résultats de l'enquête et la vérité concernant la population. Ces différences sont des distorsions causées par la non-observation, et elles peuvent être graves.

4. Bien entendu, il se peut que la distorsion due à la non-observation ne se produise pas, même lorsqu'il n'est pas effectué de mesures sur une partie de la population. S'il est assez facile de noter les cas de non-observation, la détection des distorsions qu'elle peut engendrer est difficile. C'est cette difficulté qui explique que l'étude de cette distorsion sert rarement de thème de recherche. On peut trouver des cas où la non-observation ne fait aucune différence dans les résultats de l'enquête, ou dans les réponses à la plupart des questions. Dans d'autres cas, la non-observation cause une forte distorsion dans l'estimation tirée d'une seule question ou dans les estimations tirées d'un ensemble de questions, et ce sont tous les résultats de l'enquête qui deviennent suspects.

5. La non-observation a donné lieu à une grande quantité de travaux de recherche. Ce chapitre n'offre qu'un aperçu de la nature des erreurs dues à la non-couverture et à la non-réponse dans les enquêtes sur les ménages. Le lecteur est prié de se reporter aux références pour un examen plus approfondi de ces questions. La section suivante offre un cadre destiné à permettre de distinguer la non-couverture de la non-réponse. Elle est elle-même suivie de sections distinctes sur chacune de ces sources d'erreur.

B. Cadre pour la compréhension de l'erreur de non-couverture et de l'erreur de non-réponse

6. Pour voir la différence entre la non-couverture et la non-réponse, il faut comprendre la nature des populations et des cadres d'échantillonnage. La population cible est l'ensemble des éléments sur lesquels le concepteur de l'enquête cherche à produire des estimations. Par exemple, le concepteur peut être invité à concevoir une enquête pour étudier la participation à la main-d'œuvre des personnes âgées de 15 ans ou plus vivant dans un pays donné. Cette population a des limites géographiques bien définies (les frontières du pays) et des limites quant aux caractéristiques des sujets de l'enquête (l'âge).

7. La définition de la population cible comporte d'autres considérations implicites, par exemple, ce que signifie une personne vivant dans le pays. De nombreuses enquêtes appliquent une définition de la résidence selon laquelle une personne doit avoir vécu dans le pays la majeure partie de l'année précédente ou, s'y étant installée récemment, a l'intention d'y vivre en permanence. Certaines fractions de la population peuvent déborder le cadre d'un certain thème de l'enquête. Par exemple, les personnes vivant en détention ou appartenant à des institutions telles que les forces armées peuvent se définir comme n'entrant pas dans le cadre de certaines enquêtes sur les conditions économiques. Ainsi, ces institutions peuvent être exclues car les

personnes qu'elles regroupent ne sont pas incluses dans la base conceptuelle de la mesure à effectuer. La définition de la population cible présente aussi implicitement une dimension temporelle. L'enquête s'intéresse probablement à la participation actuelle à la main-d'œuvre et non pas aux tendances du comportement des individus dans le temps. Dans ce cas, elle vise à faire des estimations des caractéristiques de la population telle qu'elle se présente à un moment particulier.

8. La population cible est aussi une population de référence. En fin de compte, les résultats de l'enquête sont censés se référer à une population particulière. Les enquêtes sont souvent conçues pour mesurer les caractéristiques des habitants d'un pays donné. Que certaines personnes soient ou non couvertes par le processus d'échantillonnage, le rapport final de l'enquête peut comporter des déclarations appliquées sans réserves à l'ensemble de la population. Par exemple, même si les personnes vivant en institution n'ont pas été incluses dans l'enquête, il se peut que le rapport final affirme que les résultats de celle-ci soient présentés comme s'appliquant à l'ensemble des personnes vivant dans le pays. Le lecteur non informé peut supposer que ces résultats concernent aussi les personnes vivant dans des institutions, même si ces dernières ont été exclues du processus d'échantillonnage. C'est pourquoi il importe, dans les publications où est présentée l'enquête, d'indiquer soigneusement et sans rien omettre quelle a été la population cible.

9. La population cible diffère souvent d'une autre population importante, à savoir l'ensemble des éléments à partir desquels l'échantillon est sélectionné, c'est-à-dire les éléments figurant dans le cadre d'échantillonnage. Le cadre d'échantillonnage est l'ensemble des documents utilisés pour la sélection de l'échantillon, et il peut arriver qu'il ne coïncide pas exactement avec la population cible. Par exemple, dans certains pays, ce sont les registres d'adresses préparés et tenus à jour par une institution de sûreté publique, telle que la police, qui servent de cadre d'échantillonnage. Or, certaines familles peuvent ne pas figurer dans ces registres administratifs. Dans ce cas, le cadre diffère de la population cible.

10. Dans d'autres cas, le cadre diffère de la population cible pour des raisons structurelles ou délibérées. Une partie de la population peut être laissée en dehors du cadre pour des raisons administratives ou des raisons de coût. Par exemple, il se peut qu'une région, plusieurs districts ou une province du pays soient le théâtre de troubles civils. Les institutions de sûreté publique peuvent alors imposer des restrictions à l'entrée ou à la sortie de ces territoires. Le concepteur de l'enquête peut décider de les exclure délibérément du cadre, même s'il existe des documents à partir desquels tirer un échantillon de leur population.

11. La décision d'exclure une partie de la population peut être motivée par des considérations de coût. Dans beaucoup de pays, les personnes vivant dans des régions reculées ou à population éparse sont exclues du cadre d'échantillonnage en raison du coût élevé que comporterait l'enquête si elles étaient sélectionnées. En outre, dans les pays où sont parlées de nombreuses langues vernaculaires, la traduction et la nécessité de recruter des enquêteurs qui puissent parler toutes ces langues sont coûteuses, et les concepteurs de l'enquête, de concert avec ses promoteurs, peuvent choisir d'exclure systématiquement les personnes qui ne parlent pas l'une des principales langues du pays. En pareil cas, on ne peut exclure une personne qu'après que la famille a été identifiée et que les aptitudes linguistiques de ses membres ont été déterminées. L'exclusion se fait par filtrage à l'intérieur des ménages.

12. Par ailleurs, les concepteurs de l'enquête peuvent décider de classer ce type de problème comme non-réponse, c'est-à-dire non-couverture pour raison de langue ou non-réponse due à l'impossibilité de communiquer. La décision quant à la façon de

classer les « exclusions pour raison de langue » dépend en partie de l'ampleur du problème. Par exemple, dans un pays, l'enquête peut être limitée aux populations qui parlent l'une des multiples langues officiellement reconnues. Cette décision risque d'exclure d'importantes fractions de la population qui ne parlent aucune de ces langues. En revanche, dans un autre pays où presque tout le monde parle l'une des langues officielles, les petits groupes de population qui parlent une langue non officielle dans laquelle le questionnaire n'a pas été traduit peuvent être contactés mais pas interrogés. Dans le premier cas, il peut être bon, avec une documentation minutieuse à l'appui, de classer les groupes exclus pour raison de langue comme cas de non-couverture. Dans le second, il suffit de classer les groupes n'ayant pas été interrogés comme cas de non-réponse.

13. Il y a non-couverture lorsque des éléments de la population cible ne correspondent pas aux listes du cadre d'échantillonnage. Dans les enquêtes sur les ménages, on est en présence d'un cas typique de non-couverture lorsque des unités d'habitation ne sont pas incluses dans la liste préparée lors des opérations sur le terrain, lorsque l'on utilise des listes administratives des ménages périmées ou inexactes ou lorsque des membres d'une famille sont omis de la liste des membres résidents de cette famille.

14. La non-couverture se produit lorsque l'on prive un élément de la population de la possibilité d'être sélectionné pour l'échantillon de l'enquête, tandis que la non-réponse correspond à une tentative infructueuse de collecter des données auprès d'une unité sélectionnée de la population-cible. La non-couverture est due à des erreurs ou à des problèmes liés au cadre utilisé pour la sélection de l'échantillon; la non-réponse se produit une fois que les cadres ont été construits et que des éléments de ce cadre ont été sélectionnés. Par exemple, supposons que dans un ménage sélectionné, un élément mâle soit absent au moment de l'entrevue parce qu'il passe la semaine en déplacement pour un travail temporaire hors du village où habite le ménage. Si cet élément ne figure pas sur la liste donnant la composition des ménages lors de l'entrevue initiale parce que le membre interrogé oublie de le mentionner, on est en présence d'un cas de non-couverture. En revanche, si un élément figure sur cette liste mais est absent au moment des entrevues effectuées dans le village et que l'enquête n'accepte que les données fournies par les éléments eux-mêmes, aucune donnée n'ayant été recueillie auprès de lui, cet élément est classé comme cas de non-réponse.

15. La non-couverture porte généralement sur des unités entières, telles que ménages ou personnes. La non-réponse peut porter sur des unités entières, ou sur certaines données particulières. Par exemple, la non-couverture peut porter sur l'absence d'un ménage d'une liste établie sur son village, parce que ce ménage habite au-dessus d'une boutique de détail. L'unité tout entière est alors absente du cadre. La non-réponse peut être due au fait qu'un ménage figurant sur la liste refuse de participer à l'enquête ou que certains de ses membres acceptent de coopérer tandis que d'autres sont absents ou refusent simplement de répondre aux questions. Ces deux formes de non-réponse d'une unité, d'un ménage ou de l'un de ses membres, diffèrent du cas où un membre d'un ménage donne des réponses à toutes les questions de l'enquête sauf à l'une des sections du questionnaire. Par exemple, un membre du ménage interrogé peut refuser de fournir des données sur le revenu qu'il ou elle tire de l'économie informelle par crainte de représailles administratives à un revenu non déclaré. Ce cas de non-réponse est appelé non-réponse ponctuelle. Il convient de noter que ce type de non-réponse varie également selon que l'unité de l'analyse est la personne ou le ménage : la non-réponse d'une personne est un cas de non-réponse ponctuelle pour l'analyse au niveau des ménages, mais un cas de non-réponse unitaire pour l'analyse au niveau des personnes.

16. Il importe également de tenir compte des arbitrages entre la non-couverture et la non-réponse. Si l'on peut identifier de nombreuses sources de non-couverture et de non-réponse pour une enquête donnée par une étude minutieuse, et s'il est vrai que l'on peut chercher à réduire l'ampleur de l'un ou l'autre de ces problèmes, cette réduction nécessite que l'on dépense une part des ressources maigres et limitées affectées à l'enquête. Il peut y avoir concurrence entre ces deux sources d'erreur pour l'affectation de ces ressources.

17. Par exemple, supposons que, dans un pays qui compte 40 langues ou dialectes importants, l'instrument de l'enquête ait été traduit dans cinq langues parlées par des ménages représentant 80 % de la population. La sixième langue par ordre d'importance est elle-même parlée par 3 % de la population. En même temps, supposons que l'enquête spécifie deux visites par ménage sur une période de deux jours, afin de trouver quelqu'un à la maison, et que l'on sache que 10 % des ménages visités deux fois ne répondront pas car personne ne se trouve à la maison durant les deux jours prévus pour les entrevues. Le concepteur de l'enquête a un choix à faire quant à l'utilisation de ses ressources. Il peut décider de consacrer plus de fonds à traduire l'instrument dans la sixième langue pour couvrir les 3 % supplémentaires de la population qui parlent cette sixième langue, ou utiliser ces fonds pour permettre aux enquêteurs de passer un troisième ou un quatrième jour dans chaque village afin d'essayer de trouver une plus forte proportion des membres des ménages chez eux.

18. La décision d'utiliser des fonds supplémentaires pour la traduction ou pour des visites additionnelles aux ménages sera fonction des inclinations prévues et des coûts et ressources en jeu. L'inclination dépend à la fois du niveau de non-couverture ou de non-réponse et des différences entre les populations couvertes et non couvertes, et les personnes de l'échantillon qui répondent et de celles qui ne répondent pas.

19. Ces types d'arbitrage entre coût et erreur sont fréquents dans la conception des enquêtes. Il déborde le cadre de ce chapitre d'examiner de façon détaillée le type de données nécessaires pour décider dans un sens ou dans l'autre et la façon dont ces décisions sont prises. Pour la plupart des enquêtes, ces arbitrages se font sur la base d'informations limitées ou de façon informelle.

C. Erreur de non-couverture

1. Sources de la non-couverture

20. Les sources de la non-couverture dans les enquêtes sur les ménages sont fonction des informations contenues dans le cadre qui sont utilisées pour choisir l'échantillon. Comme beaucoup d'enquêtes sur les ménages effectuées dans les pays en développement et dans certains pays en transition utilisent des méthodes d'échantillonnage géographique, la présente analyse se limitera aux problèmes de cadre et de non-couverture que posent les enquêtes sur les ménages fondées sur des échantillons géographiques.

21. L'échantillonnage géographique est souvent combiné à une sélection à phases multiples. Les phases primaires et parfois secondaires de la sélection portent sur des zones géographiques qui peuvent être considérées comme des grappes de ménages. À la phase suivante de la sélection, il faut se procurer ou créer une liste des ménages couvrant un ensemble relativement restreint de zones géographiques. À la dernière phase de la sélection, on établit une liste des personnes ou résidents d'un ménage pour chaque zone géographique. Il y a donc trois types d'unités à prendre en compte lorsque l'on examine la non-couverture de ces enquêtes : les unités géographiques, les

ménages et les personnes. Comme on le verra plus avant, ces unités peuvent aussi être chacune source de non-réponse dans les enquêtes sur les ménages.

22. La non-couverture des zones géographiques par suite de déficiences du cadre d'échantillonnage est rare, parce que la plupart des cadres géographiques sont fondés sur des données de recensement qui couvrent toute l'étendue géographique habitée par la population. La non-couverture d'une zone géographique peut effectivement se produire, mais sous une forme plus subtile, comme on l'a indiqué précédemment. Une enquête peut être conçue pour permettre de tirer des conclusions applicables à l'ensemble de la population du pays ou de l'une de ses régions, et les références à la population figurant dans le rapport final peuvent en fait s'appliquer à la population du pays tout entier alors que l'échantillon n'a pas été sélectionné sur tout le territoire du pays en question.

23. Par exemple, lors de la conception de l'enquête, les concepteurs peuvent identifier certaines zones géographiques qui ne comptent qu'une petite fraction de la population et qui seraient extrêmement coûteuses à couvrir. Ils peuvent décider délibérément d'exclure ces zones géographiques du cadre. Cependant, en rendant compte des résultats de l'enquête, ils ne mentionnent pas, ou ne mentionnent que brièvement, l'exclusion de ces zones. Les lecteurs peuvent avoir l'impression, ou être amenés implicitement à croire, que les résultats de l'enquête s'appliquent à la population de l'ensemble du pays ou de la région, alors qu'en fait une partie de cette population n'est pas couverte. Dans la pratique, l'ampleur de l'erreur de non-couverture découlant de telles situations est généralement faible, et le plus souvent ignorée.

24. Il ne faut pas perdre de vue que la distinction reste entre une population cible désirée (en l'occurrence, la population vivant sur l'ensemble du territoire géographique du pays) et une « population étudiée » restreinte, vivant dans la zone géographique incluse dans l'enquête. On court toutefois le risque que, du fait d'une documentation incomplète, l'usager des données s'imagine que l'échantillon de l'enquête couvre la population tout entière alors qu'en fait il n'en est rien.

25. Une source de non-couverture plus importante se situe au niveau des ménages. La plupart des enquêtes considèrent les ménages comme l'ensemble des personnes vivant habituellement dans une même unité d'habitation. Deux composantes sont donc importantes : la définition du résident habituel et celle de l'unité d'habitation.

26. Les définitions des unités d'habitation sont complexes, car elles doivent tenir compte de la question de savoir si une construction est destinée à être utilisée comme logement, et si les personnes qui y habitent y vivent et y prennent leurs repas séparément d'autres personnes habitant dans la même construction (comme dans les constructions à logement multiples telles que les immeubles à appartements). Le fait de vivre séparément implique que les résidents ont directement accès à leur logement depuis l'extérieur de la construction ou depuis un hall d'entrée commun. L'aptitude à « prendre les repas séparément » implique généralement l'existence d'un endroit où préparer les repas, et l'entière liberté des résidents de choisir ce qu'ils veulent manger.

27. Il est difficile d'appliquer ce type de définition générale aux nombreuses situations diverses qui peuvent se présenter à travers les pays, ou à travers les régions d'un pays. La plupart des unités d'habitation sont faciles à identifier, par exemple, maisons unifamiliales ou individuelles, maisons jumelées à logements séparés par un mur mitoyen, avec entrées séparées, et appartements d'immeubles. Cependant, de nombreuses unités d'habitation sont difficiles à classer ou à trouver. Par exemple, dans les bidonvilles urbains, les unités d'habitation sont difficiles à discerner lorsque leurs habitants vivent dans des constructions faites à partir de matériaux recyclés ou de maté-

riaux de récupération. Les unités d'habitation peuvent être situées dans des endroits qui ne peuvent être repérés par une simple inspection des entrées depuis une rue, une allée ou un chemin.

28. Dans les zones rurales, une construction destinée à servir de logement est facile à identifier, mais il peut être difficile d'y distinguer divers logements par suite d'arrangements sociaux complexes. Par exemple, dans les groupes tribaux, le logement se fait dans de longues bâtisses à entrée unique, avec plusieurs cloisons à l'intérieur pour permettre à plusieurs familles d'y dormir, mais avec un seul endroit pour la préparation des repas pour le groupe ou pour chaque famille individuellement, de sorte que les compartiments formés par les cloisons ne forment pas d'eux-mêmes des unités d'habitation car l'entrée est commune, de même que la cuisine et le réfectoire. Dans ce cas, la notion de ménage en tant que groupe de personnes vivant généralement dans une unité d'habitation clairement définie est plus difficile à appliquer. On ne sait pas très bien si la bâtisse tout entière, ou chacun de ses compartiments, doit être traité comme unité d'habitation. Dans la pratique, on traite comme telle la bâtisse toute entière, et si elle est sélectionnée pour l'échantillon, tous les ménages identifiés lors de l'établissement des listes de ménages sur le terrain sont inclus dans l'enquête.

29. Il y a aussi des locaux d'habitation qui ne sont pas considérés comme des unités d'habitation. Les bâtiments où vivent des personnes placées sous la garde d'autres personnes, comme les orphelinats, les prisons ou les hôpitaux, ne sont pas considérés comme unités d'habitation. Les foyers d'étudiants, les monastères et les couvents et les abris pour sans-logis sont des types spéciaux de locaux d'habitation qui n'assurent pas nécessairement la garde ou les soins attendus d'une institution. Les foyers pour gens de passage ou travailleurs saisonniers sont un autre cas difficile à traiter. Par exemple, il peut y avoir des unités d'habitation séparées dans une région agricole pour les ouvriers saisonniers, qui ne sont habitées que durant une ou deux saisons chaque année. Il est probable que les occupants saisonniers de ces logements vivent généralement ailleurs, et qu'ils ne devraient pas être comptés comme faisant partie de la famille à laquelle est rattachée l'unité d'habitation saisonnière.

30. L'échantillonnage géographique à plusieurs degrés dans les pays en développement nécessite qu'à un certain stade du processus d'enquête, des listes de logements soient établies pour les petites zones géographiques, telles qu'un bloc dans une ville ou une zone d'énumération en milieu rural. La non-couverture se produit souvent lorsqu'un personnel d'enquête travaillant à temps partiel est envoyé sur le terrain pour établir la liste des unités d'habitation et se heurte aux types de locaux d'habitation complexes décrits ci-dessus. L'identification de la plupart des unités d'habitation est simple, mais il arrive encore que des unités échappent à la vigilance d'un personnel à temps partiel n'ayant qu'une expérience limitée de l'application d'une définition à composantes multiples à des arrangements d'habitation complexes.

31. Le problème de la non-couverture qui se pose lors de l'établissement d'une liste d'unités d'habitation est encore compliqué par la dimension temporelle. Lors de l'établissement de la liste, il se peut qu'une unité d'habitation soit inoccupée ou en construction. Si l'enquête doit avoir lieu ultérieurement, ce type d'unité peut avoir à être inclus dans la liste. Lorsque les listes d'unités d'habitation sont utilisées pour plusieurs phases d'une même enquête ou pour plusieurs enquêtes différentes, il est fréquent d'inclure les unités inoccupées ou en construction.

32. Pour les enquêtes effectuées dans les pays en transition, il peut être possible d'utiliser une liste déjà établie par une autorité administrative. Toutefois, il convient d'évaluer soigneusement la qualité de ces listes pour les enquêtes sur les ménages. On

risque fort de se heurter également aux types de problèmes qui peuvent se poser avec les listes établies pour les enquêtes lorsque l'on utilise les listes administratives.

33. Ainsi, le processus d'établissement de listes d'unités d'habitation peut être cause de la non-couverture de certains types de logements. Cette non-couverture peut être difficile à identifier sans l'investissement d'un volume substantiel de ressources additionnelles.

34. Enfin, à l'intérieur d'une unité sélectionnée, l'établissement de la liste des personnes qui vivent habituellement dans cette unité fait partie du processus d'établissement de la liste des ménages. Il faut des règles indiquant aux enquêteurs qui inclure dans une unité d'habitation comme résident habituel. Comme dans le cas des unités d'habitation, la détermination est le plus souvent simple à faire. La plupart des personnes rencontrées habitent dans cette unité au moment où celle-ci est contactée, et cette unité est leur unique lieu de résidence. Il y a d'autres personnes qui sont absentes au moment de ce contact, mais dont l'unité d'habitation est l'unique lieu de résidence.

35. Toutefois, il y a des personnes qui vivent dans plusieurs unités d'habitation. Il appartient au personnel à temps partiel travaillant sur le terrain de décider si les unités d'habitation sélectionnées sont ou non le lieu de résidence habituelle de ces personnes. Il peut être également difficile aux membres du ménage interrogés de rendre compte exactement des arrangements de certains résidents. Ces informations sur un autre résident ne sont pas toujours totalement exactes.

36. Les déclarants peuvent aussi avoir des raisons personnelles de refuser délibérément d'admettre que certains membres habitent avec eux. Par exemple, le fait qu'une personne habite dans l'unité d'habitation pourrait faire perdre au ménage le droit à certaines prestations qu'il reçoit de l'État. De plus, le déclarant peut exclure sciemment un résident qui ne tient pas à être identifié par des organismes publics ou privés en raison de difficultés financières (par exemple, de dettes) ou de problèmes avec la justice (en raison d'une activité criminelle).

37. Les déclarants peuvent également choisir de ne pas inclure quelqu'un comme membre de la famille pour des raisons culturelles. Ils peuvent décider de ne pas déclarer un enfant de moins d'un an parce que leur culture ne le considère pas comme assez âgé pour être traité comme une personne. Ils peuvent également exclure ces enfants parce qu'ils estiment que l'organisme chargé de l'enquête ne tient pas à recueillir des données sur les jeunes enfants, ou ils peuvent simplement oublier de compter quelqu'un comme membre de la famille, qu'il s'agisse d'un jeune enfant ou d'une personne plus âgée.

38. La non-couverture dans les enquêtes sur les ménages peut donc découler de toute une variété de conditions de définition et de conditions opérationnelles. La question est de savoir dans quelle mesure elle entraîne une erreur dans les résultats de l'enquête.

2. Erreur de non-couverture

39. Supposons que l'enquête cherche à estimer la moyenne d'une certaine caractéristique Y pour une population de N personnes, dont N_{nc} ne sont pas couvertes par le cadre d'échantillonnage de l'enquête. Supposons que la moyenne de la population de taille N est \bar{Y} , que \bar{Y}_c est la moyenne des personnes couvertes par le cadre d'échantillonnage, et que \bar{Y}_{nc} est la moyenne des personnes non couvertes par ce cadre. L'erreur liée à la non-couverture est appelée distorsion liée à la non-couverture de la moyenne de l'échantillon, \bar{Y}_c , qui est fondée uniquement sur les personnes couvertes par l'échantillon, et qui en fait donne une estimation de \bar{Y}_c plutôt que de \bar{Y} .

40. La distorsion de la moyenne de l'échantillon, \bar{Y}_c , dépend de deux composantes, la proportion de la population qui n'est pas couverte, N_{nc}/N , et la différence de moyenne de la caractéristique Y entre les personnes couvertes et les personnes non couvertes. Autrement dit :

$$B(\bar{y}_c) = (N_{nc} / N)(\bar{Y}_c - \bar{Y}_{nc})$$

41. Cette formulation de la distorsion liée à la non-couverture aide à comprendre comment les concepteurs d'enquêtes font face à la non-couverture. Pour maintenir faible l'erreur liée à la non-couverture, ou pour réduire son effet, le concepteur de l'enquête doit soit avoir de légères différences entre les personnes couvertes et non couvertes, soit avoir une faible proportion des personnes non couvertes par l'enquête.

42. La difficulté importante que comporte cette formulation est que, dans la plupart des enquêtes, on ne connaît ni la différence ($\bar{Y}_c - \bar{Y}_{nc}$) ni la proportion (N_{nc}/N) de personnes non couvertes. De plus, il peut également arriver que le taux de non-couverture (N_{nc}/N) varie entre les sous-catégories. La différence peut aussi varier selon les variables et selon les sous-catégories de personnes (telles que régions ou sous-groupes définis par une caractéristique démographique, telle que l'âge). Ainsi, l'erreur liée à la non-couverture est une propriété non pas de l'enquête mais de la caractéristique particulière, et de la statistique estimée.

43. Dans beaucoup d'organisations gouvernementales d'enquête, on exige fréquemment des estimations du total. La distorsion liée à la non-couverture associée à un total dépend non seulement des différences entre les unités couvertes et non couvertes dans la caractéristique pertinente, mais aussi du nombre (et non pas du taux) des unités non couvertes; cela revient à dire que, avec un total estimé de personnes interrogées $\hat{Y}_r = N \bar{y}_r$, la distorsion est $B(\hat{Y}_r) = N_{nc} (\bar{Y}_r - \bar{Y}_m)$.

Réduction, mesure et notification de l'erreur de non-couverture

44. Il y a quatre moyens possibles de traiter l'erreur de non-couverture dans les enquêtes sur les ménages :

- Réduire le niveau de non-couverture par l'amélioration des procédures employées sur le terrain
- Créer des procédures pour mesurer l'ampleur de l'erreur de non-couverture et en indiquer le niveau dans l'enquête.
- Tenter de compenser l'erreur de non-couverture par des ajustements statistiques.
- Indiquer les propriétés de non-couverture de l'enquête de façon aussi complète que possible dans le rapport de l'enquête.

45. On cherche généralement à réduire l'erreur de non-couverture dans les enquêtes sur les ménages en utilisant des cadres multiples ou en employant des méthodes visant à améliorer les processus d'établissement de listes. Les cadres multiples sont plus couramment utilisés pour les unités d'habitation que pour les personnes. Ils supposent l'existence de listes d'unités d'habitations distinctes qui posent des problèmes particuliers pour l'établissement de listes sur le terrain.

46. Par exemple, supposons qu'il soit difficile d'établir sur le terrain des listes d'unités d'habitation saisonnières pour ouvriers agricoles dans un pays donné. Supposons aussi que l'administration responsable de la production agricole, de l'éducation ou du bien-être social ait une liste du nombre et des types d'unités d'habitation saisonnières en place dans les exploitations ou les entreprises qui emploient et logent une

main-d'œuvre saisonnière. Une liste d'unités d'habitation saisonnières émanant d'une autre source peut être utilisée comme cadre distinct. Les enquêteurs chargés de préparer des listes d'unités d'habitation sur le terrain recevraient une liste des exploitations ou des entreprises où les listes de l'administration de tutelle seraient déjà disponibles et seraient dispensés d'établir des listes des unités d'habitation saisonnières correspondantes. Des échantillons d'unités d'habitation seraient alors choisis pour l'enquête à partir de la liste d'unités d'habitations préparée par l'enquêteur et de la liste tenue par l'administration. Ces deux listes laisseraient vraisemblablement des unités non couvertes, de même qu'elles pourraient présenter un certain degré de « sur-couverture », mais l'utilisation des deux cadres tendrait à réduire le niveau de non-couverture et l'erreur qui en découle.

47. Il importe également d'envisager des moyens d'améliorer les processus d'établissement de ces listes. Lorsqu'une administration dispose de listes d'unités d'habitation, ces listes peuvent faire l'objet d'une actualisation sur le terrain avant qu'un échantillon ne soit prélevé. Les enquêteurs peuvent être envoyés dans des zones géographiques munis d'une liste d'unités d'habitation de cette administration et d'instructions sur la façon de vérifier cette liste et d'y ajouter ou d'en retrancher des unités d'habitations selon leurs constatations.

48. Les enquêteurs peuvent également être initiés à l'utilisation sur le terrain de la procédure « de l'intervalle à demi-ouvert » pour repérer les unités d'habitation qui ne figurent pas sur les listes de l'administration ou sur les listes établies sur le terrain. Cette procédure comprend la sélection d'une unité d'habitation à partir d'une liste d'adresses, la visite d'un enquêteur à l'unité sélectionnée et une vérification implicite ou explicite de la liste. Sur place, l'enquêteur est chargé de se renseigner sur toutes unités d'habitation supplémentaires qui pourraient se trouver entre l'unité sélectionnée et la suivante sur la liste.

49. L'unité suivante sur la liste est définie à partir d'un itinéraire prédéterminé à travers la zone géographique. Par exemple, à la ville, dans un bloc d'habitations, les enquêteurs chargés d'établir la liste sont censés commencer à un angle donné, puis progresser dans le sens des aiguilles d'une montre autour de ce bloc. La liste des unités d'habitation de ce bloc est censée être dressée par énumération de ces unités prises dans le sens des aiguilles d'une montre.

50. Si l'enquêteur s'aperçoit qu'une unité d'habitation ne figure pas sur la liste entre l'unité sélectionnée et la suivante, il lui appartient d'ajouter cette unité manquante à l'échantillon et de tenter d'y obtenir une entrevue. Si plusieurs unités sont manquantes, l'enquêteur peut avoir à solliciter de nouvelles instructions auprès de l'office central chargé de l'enquête, afin d'éviter des difficultés dans les opérations sur le terrain.

51. Au sein des ménages, l'amélioration des listes peut nécessiter que l'enquêteur pose une série de questions au déclarant de l'unité d'habitation afin d'identifier les personnes qu'il n'a pu rencontrer. Par exemple, il peut avoir à poser des questions au sujet de jeunes enfants qui ont pu être exclus de la liste des occupants habituels de l'unité. On peut également améliorer la liste des ménages en donnant des directives aux enquêteurs sur le choix optimal des déclarants ou en leur demandant de répéter les noms figurant sur la liste des personnes devant le déclarant afin de s'assurer que personne n'a été omis.

52. La mesure de la distorsion due à la non-couverture est aussi un facteur important, quoique difficile à traiter. Comment l'organisme chargé de l'enquête peut-il identifier les unités qui ne figurent sur aucune liste ? La mesure de la non-couverture peut être une opération coûteuse, et elle n'est effectuée qu'occasionnellement.

53. Le moyen courant d'évaluer l'erreur de non-couverture est de comparer les résultats de diverses enquêtes et d'examiner les variables pour lesquelles on peut faire des comparaisons avec les conclusions obtenues de sources externes ou indépendantes. Pour mesurer l'ampleur de la non-couverture, l'enquête peut chercher à comparer la répartition par âge et par sexe des personnes de l'échantillon à la répartition qui se dégageait d'un recensement récent ou des fichiers administratifs. Les différences de répartition sont révélatrices de problèmes de non-couverture. Pour mesurer l'erreur de non-couverture liée à une variable, on peut comparer les grandeurs de la statistique étudiée avec celles obtenues d'une source indépendante. Par exemple, le total des revenus salariaux signalé par une enquête pour l'ensemble de l'échantillon et pour certains sous-groupes clés peut être comparé aux données des rapports administratifs sur ces revenus. Dans une étude classique, Kish et Hess (1950) ont comparé la répartition des unités d'habitation d'une enquête aux données d'un recensement récent sur la répartition de ces unités par bloc. Cette comparaison a permis de se faire une idée de la nature du problème de non-couverture qui s'est posé lors de la collecte de données pour l'enquête.

54. On peut procéder à une évaluation plus coûteuse de l'erreur de non-couverture au moyen d'un double système de mesures, ou de procédures de comparaison avec des cas apparentés. Les recensements emploient le double système pour évaluer la couverture d'un recensement [voir, par exemple, Marks (1978)]. Lors d'un recensement, on compare les données d'une enquête distincte aux résultats du recensement afin d'identifier les problèmes de non-couverture. L'évaluation de l'ampleur de ce problème passe par une comparaison cas par cas de l'échantillon de l'enquête aux éléments du recensement afin de déterminer les éléments de l'échantillon qui n'apparaissent pas dans le recensement. Ces procédures sont très proches des méthodes « d'échantillonnage par saisie-resaisie » utilisées dans les études environnementales des populations animales.

55. Comme les enquêtes sur les ménages sont toutes affectées pour l'erreur de non-couverture, de nombreuses enquêtes emploient les ajustements post-stratification ou ajustements de contrôle de la population comme procédures statistiques pour ajuster les résultats de l'enquête de manière à compenser l'erreur de non-couverture. Ces ajustements sont très comparables à la méthode décrite ci-dessus employée pour évaluer l'ampleur de l'erreur de non-couverture. Par exemple, on peut comparer la distribution de l'échantillon par âge et par sexe à celle émanant d'une source extérieure, telle qu'un recensement récent ou de récentes projections de la population. Lorsque la distribution est faible (ou élevée) pour un groupe d'âge et de sexe, on peut appliquer une pondération à toutes les données sur les personnes de l'échantillon de ce groupe pour accroître (diminuer) leur contribution aux résultats de l'enquête. Il faut néanmoins des estimations pondérées pour utiliser comme il convient les pondérations dans cette analyse.

56. La dernière remarque concernant la non-couverture est que la préparation de rapports de qualité est importante pour tout organisme statistique. Les rapports analytiques doivent donner des définitions claires de la population cible, et mentionner notamment les exclusions. Le cadre doit être décrit de façon suffisamment détaillée pour que le lecteur puisse constater comment la non-couverture pourrait se produire, ou même procéder à une évaluation informelle de l'ampleur potentielle de l'erreur. Il serait utile d'inclure comme références ou annexes toutes évaluations de la qualité du cadre, tels que contrôles de la qualité des listes d'unités d'habitation ou des listes administratives, ou comparaisons des listes originales de personnes vivant dans les unités d'habitation aux listes obtenues à la suite de nouvelles entrevues effectuées aux fins de contrôle de qualité.

57. Il est plus difficile de rendre compte des taux de non-couverture ou de la distorsion due à la non-couverture pour la population ou les sous-catégories de population. Ces types d'évaluations ne sont possibles que pour les enquêtes en cours durant lesquelles, à un certain stade, on a cherché à mesurer l'ampleur du problème de la non-couverture. Il est très difficile, sinon impossible, d'effectuer de telles évaluations pour des enquêtes transversales ponctuelles.

58. Enfin, s'il est procédé à une post-stratification ou à des ajustements de contrôle de la population, la documentation de l'enquête doit comprendre une description des procédures d'ajustement et de l'ampleur de ceux-ci pour les sous-groupes de population importants.

D. Erreur de non-réponse

59. L'erreur de non-réponse suggère un certain nombre de parallèles avec l'erreur de non-couverture en ce qui concerne les définitions, la mesure, la réduction, la compensation et la notification. La présente section est donc articulée d'une façon très comparable à la section C. Il importe toutefois de préciser que la non-réponse et la non-couverture sont deux problèmes très distincts, dont les sources et, dans certains cas, les solutions sont différentes. Alors que pour la non-couverture, les concepteurs de l'enquête ne savent presque jamais rien d'autre que l'endroit et les caractéristiques générales de la partie de la population non couverte, pour la non-réponse, ils disposent au moins des informations du cadre pour les non-répondants. On estime également que la non-réponse est plus répandue dans les enquêtes sur les ménages, et que sa contribution aux distorsions des estimations découlant de ces enquêtes est probablement plus forte.

60. Comme on l'a indiqué précédemment, on constate souvent deux types de non-réponse dans les enquêtes sur les ménages, à savoir la non-réponse unitaire et la non-réponse ponctuelle. Ces deux types ont des conséquences très différentes pour les résultats de l'enquête, et les méthodes employées pour les mesurer, pour les réduire et pour en rendre compte et les compenser sont également différentes à certains égards. S'il est vrai que l'on pourrait consacrer une section à chacun de ces deux types, ceux-ci seront examinés ensemble dans la présente section.

1. Sources de non-réponse dans les enquêtes sur les ménages

61. Dans les enquêtes sur les ménages, les cas de non-réponse unitaire peuvent se produire pour différents types d'unités. De même que pour la non-couverture, la non-réponse peut se produire pour des unités d'échantillonnage primaires ou secondaires. Par exemple, une unité d'échantillonnage primaire peut consister en un district ou un sous-district d'un pays. Les conditions climatiques ou une catastrophe naturelle peuvent empêcher l'enquête de se dérouler dans un district ou un sous-district qui a été choisi à un stade d'échantillonnage primaire ou secondaire. L'unité est couverte par l'enquête mais pendant que celle-ci se déroule, il n'est possible de recueillir des données auprès d'aucun ménage de cette unité.

62. La non-réponse est plus fréquente au niveau des ménages. Il peut arriver qu'une unité d'habitation figurant sur une liste et choisie pour l'échantillon soit trouvée occupée et qu'une entrevue y soit tentée. Toutefois, lorsque l'enquêteur se rend dans cette unité, une combinaison de facteurs peut très bien l'empêcher de procéder à la collecte de données. L'un des membres du ménage peut refuser sa participation, que ce soit comme individu ou comme représentant de l'unité dans son ensemble.

63. Bien qu'une unité d'habitation soit occupée, il peut arriver que ses habitants soient absents de chez eux pendant toute la durée de l'enquête. Dans certains pays en développement, le fait que certaines unités d'habitation sont clairement habitées, mais inaccessibles pendant toute la période de collecte de données, pose un réel problème.

64. Dans beaucoup de pays, bien que des unités d'habitation occupées aient des personnes présentes au moment de la collecte de données, la langue peut être un véritable obstacle. Le questionnaire de l'enquête n'a peut-être pas été traduit dans la langue du ménage, ou l'enquêteur ne parle peut-être pas la langue locale. Pour éviter la non-réponse, des interprètes peuvent être recrutés sur place pour accompagner les enquêteurs lors de leurs entrevues. D'autres enquêtes rejettent cette pratique en raison de doutes sur l'exactitude des traductions et sur leur cohérence d'un ménage à un autre. Dès lors, les ménages qui ne peuvent fournir de réponses risquent d'être classés comme unités non répondantes. L'autre solution pratiquée consiste à exclure de l'enquête les ménages qui ne parlent pas une langue de traduction. Ces ménages sont alors considérés comme cas de non-couverture plutôt que de non-réponse. L'approche choisie par l'organisme enquêteur, à savoir de traiter ces unités comme non couvertes ou non répondantes, doit être clairement décrite dans la documentation de l'enquête.

65. La non-réponse peut également se produire au niveau des personnes. Pour les enquêtes qui autorisent les réponses par quelqu'un d'autre aux questions de l'enquête, les données peuvent être recueillies auprès d'autres membres de la famille au nom des personnes qui ne sont pas là au moment de l'entrevue. En revanche, pour les enquêtes qui exigent une réponse personnelle à certaines ou à l'ensemble des questions, une personne qui est absente au moment de l'enquête, qui refuse de participer ou qui se heurte à une autre difficulté (par exemple, linguistique) qui l'empêche de répondre, est considérée comme non répondante. Un handicap de santé, qu'il soit permanent comme la surdité ou la cécité, ou passager comme une maladie, peut aussi empêcher un individu de répondre.

66. Au même titre que les ménages ayant des problèmes de langue, certains organismes d'enquête décident de classer les personnes qui se trouvent dans la même situation comme non couvertes et celles qui sont empêchées temporairement comme non répondantes (Seligson et Jutkowitz, 1994). Il n'y a pas de règles générales qui permettent de décider de la façon de faire cette classification. Dans le cas d'une enquête sur les revenus ou les dépenses, les personnes souffrant de problèmes de santé passagers sont assez peu nombreuses pour que l'organisme puisse les traiter comme non couvertes. En revanche, s'il s'agit d'une enquête sur l'état de santé, les réponses de ces individus peuvent s'écarter suffisamment de celles des autres pour que l'on hésite à les exclure. Ces personnes peuvent alors être classées comme non répondantes. Compte tenu de l'absence de pratique généralement acceptée, il importe que les organismes d'enquête indiquent clairement dans leurs rapports comment ces cas ont été traités dans le cadre d'une enquête donnée.

2. Distorsion due à la non-réponse

67. Des travaux de recherche beaucoup plus abondants ont été consacrés au problème de la non-réponse dans les enquêtes sur les ménages en comparaison de celui de la non-couverture [voir, par exemple, les études de Groves et Couper (1998) et de Lessler et Kalsbeek (1992)]. Cela est lié à plusieurs facteurs.

68. Dans un certain sens, la non-couverture est moins visible que la non-réponse. Les ménages ou individus non couverts ne peuvent être étudiés, tandis que

les unités non répondantes peuvent être observées et comptées et, éventuellement, persuadées de participer.

69. Dans les pays développés, on présume que la non-couverture est moins importante que la non-réponse car le taux de non-couverture est moins élevé que le taux de non-réponse. C'est peut-être l'inverse dans les pays en développement, où les taux de non-réponse sont moins élevés et les taux de non-couverture beaucoup plus élevés que dans les pays développés. On se rappellera que la distorsion due à la non-couverture pour la moyenne d'un échantillon est imputable à deux sources : l'ampleur du taux de non-couverture et l'ampleur de la différence entre les moyennes des groupes de population couverts et non couverts. De même, pour la non-réponse, l'ampleur de la distorsion due à la non-réponse pour la moyenne d'un échantillon peut être attribuée à la proportion de la population qui ne répond pas et à l'ampleur de la différence entre les moyennes des groupes de répondants et de non-répondants.

70. Selon le même raisonnement que pour la non-couverture, supposons que l'enquête cherche à estimer la moyenne d'une certaine caractéristique Y , et que la moyenne de la population \bar{Y} soit composée d'une moyenne de personnes qui répondent, disons \bar{Y}_r , et d'une moyenne de personnes qui ne répondent pas, \bar{Y}_{nr} . Supposons que N_{nr} indique le nombre de personnes qui ne répondraient pas si elles étaient sélectionnées. La distorsion de la moyenne de l'échantillon pour les répondants \bar{Y}_r est donc $B(\bar{y}_r) = (N_{nr} / N)(\bar{Y}_r - \bar{Y}_{nr})$. Comme pour la non-couverture, le concepteur de l'enquête doit maintenir le taux de non-réponse peu élevé, ou prévoir de légères différences entre les ménages et les personnes qui répondent et ceux qui ne répondent pas. On peut utiliser ce cadre général pour comprendre également la non-réponse au niveau ponctuel. Le problème de la distorsion due à la non-réponse ponctuelle est plus compliqué, cependant, car souvent, différents points sont examinés ensemble, et la non-réponse ponctuelle est l'ensemble de non-réponses sur plusieurs points.

71. Alors que, pour la non-couverture, on ne connaît ni la différence ni le taux, pour la non-réponse, les études conçues avec soin donnent de bonnes estimations du taux de non-réponse. De telles études tiennent une comptabilité détaillée du traitement de chaque unité de l'échantillon, qu'il s'agisse d'un ménage, d'une personne ou d'un point particulier, choisis pour l'étude. Elles permettent alors d'estimer le taux de non-réponse directement à partir des données de l'enquête. Elles peuvent également observer certaines données si les taux de réponse diffèrent à l'intérieur de sous-catégories importantes, en particulier de sous-catégories géographiques pour les ménages.

72. Pour évaluer les différences entre répondants et non-répondants, il faut une collecte et une mesure plus amples des données. Lors de la collecte des données pour l'enquête, il est souvent impossible de chercher à mesurer les caractéristiques étudiées pour les non-répondants. On peut toutefois entreprendre des études spéciales parallèlement à l'enquête, afin d'obtenir des réponses des unités non répondantes.

73. La non-réponse lors des phases ultérieures d'une enquête répétée ou par panel fournit plus de données à partir desquelles étudier les effets potentiels de la distorsion due à la non-réponse que la non-réponse à une enquête ponctuelle ou transversale. Dans les enquêtes par panel, les mêmes unités sont suivies et des données sont recueillies à plusieurs reprises auprès des unités du panel. Une partie de ces unités peuvent disparaître lors du suivi, entraînant une non-réponse par attrition durant le cours de l'enquête. Toutefois, les travaux de recherche sur cette non-réponse peuvent utiliser les données collectées lors des phases précédentes pour mieux approfondir les différences entre répondants et non-répondants et pour servir de base aux types d'ajustements décrits ci-dessous. Les techniques de compensation pour la non-réponse aux enquêtes par panel sont décrites par Lepkowski (1988).

74. Le fait que l'on dispose d'un peu plus d'informations sur les non-répondants que sur les personnes non couvertes, et la possibilité d'utiliser des modèles de comportement pour étudier la non-réponse et la compenser expliquent qu'il y ait plus de travaux de recherche sur la non-réponse que sur la non-couverture. Lorsque l'on tient avec soin les dossiers sur toutes les unités de l'échantillon, et pas seulement sur celles qui répondent, on peut effectuer directement des comparaisons entre les répondants et les non-répondants à partir des données de l'échantillon. Par ailleurs, la non-réponse est engendrée en partie par le comportement du ménage ou de l'individu : c'est un phénomène d'autosélection. Le concepteur de l'enquête peut consulter d'abondants ouvrages de sociologie, psychologie et psychologie sociale pour étudier comment les individus et les groupes décident ou non de participer à diverses activités. Il peut examiner des modèles de comportement, à condition qu'il dispose de certaines données sur les non-répondants, afin de comprendre les déterminants de la non-réponse à une enquête.

3. Mesure de la distorsion due à la non-réponse

75. La mesure de la distorsion due à la non-réponse nécessite la mesure des taux de non-réponse et la mesure des différences entre répondants et non-répondants sur les variables de l'enquête. Le calcul du taux de non-réponse pour les ménages ou les personnes à partir des données de l'échantillon nécessite lui-même la définition des résultats possibles pour tous les cas sélectionnés, puis l'indication de la façon dont ces résultats devraient être utilisés pour calculer un taux. Par exemple, les réponses complètes et partielles (celles qui offrent suffisamment de données pour que l'on puisse se faire une idée sur les principaux concepts de l'étude) sont souvent groupées ensemble.

76. Les cas de non-entrevue admissibles sont ceux d'éléments qui appartiennent à la population et sont identifiés par l'enquête mais auprès desquels il n'a pas été recueilli de données. Par exemple, si une enquête est limitée aux personnes âgées de 15 ans ou plus, les cas de non-entrevue admissibles sont les personnes de ce groupe d'âge sur lesquelles il n'a pas été recueilli de données. On compte habituellement au moins trois sources de non-entrevue : les refus (Ref), c'est-à-dire les personnes ou ménages qui ont été contactés mais ont refusé de participer à l'étude; les cas de non-contact (NC), c'est-à-dire les personnes ou ménages admissibles avec lesquels le contact ne peut être établi lors de la collecte de données; et les autres (Oth), c'est-à-dire les cas de non-entrevue qui se produisent pour quelque autre raison, telle que la difficulté linguistique ou l'état de santé. Enfin, il y a aussi les cas non admissibles (Inelig) pour l'enquête (par exemple, les jeunes de moins de 15 ans) et ceux dont on ignore s'ils sont ou non admissibles (Unk).

77. Le taux de réponse pour cet ensemble simplifié de résultats peut se calculer de différentes façons. Une méthode couramment acceptée de calcul du taux de réponse (où « Int » indique le nombre d'entrevues complètes ou partielles de l'enquête) est donnée par l'équation :

$$\bar{R} = \frac{\text{Int}}{\text{Int} + \text{Ref} + \text{NC} + \text{Oth} + \varepsilon \times \text{Unk}}$$

Ici, on estime qu'une certaine proportion, ε , des cas d'admissibilité inconnue sont admissibles. Souvent, on calcule cette admissibilité estimée à partir des données existantes en utilisant le taux d'admissibilité connue (les cas avec des résultats Int,

Ref, NC et Oth) parmi tous les cas pour lesquels l'admissibilité a été déterminée. Par conséquent :

$$\hat{\varepsilon} = \frac{\text{Int} + \text{Ref} + \text{NC} + \text{Oth}}{\text{Int} + \text{Ref} + \text{NC} + \text{Oth} + \text{Inelig}}$$

78. Les enquêtes sur les ménages qui comportent des entrevues répétées avec les mêmes ménages, ou un panel de personnes sélectionnées à partir d'un échantillon de ménages, présentent d'autres considérations de non-réponse qui influent sur le calcul des taux de réponse. Ces enquêtes longitudinales par panel présentent des cas de non-réponse unitaire lors de la première vague d'entrevues comme une enquête transversale, mais en plus, elles peuvent se montrer incapables d'obtenir des données de certains membres du panel lors des vagues d'entrevues suivantes. Les calculs du taux de réponse doivent tenir compte des pertes causées par la non-réponse à la vague initiale comme aux vagues suivantes de collecte de données. Il n'entre pas dans le cadre de la présente publication de traiter le calcul des taux de réponse dans les enquêtes par panel. On trouvera plus de renseignements sur ce sujet sur le site Internet de l'American Association for Public Opinion Research (<http://www.aapor.org>. Choix : Survey Methods).

79. La mesure des différences entre les moyennes des répondants et des non-répondants, de même que d'autres statistiques, sont plus difficiles à obtenir. On peut comparer les résultats de l'enquête à ceux d'une source extérieure en ce qui concerne certaines variables, afin de déterminer s'il y a une grande différence entre l'enquête et la source extérieure pour ce qui est de la valeur d'une estimation; toutefois, cette approche peut se révéler difficile à appliquer car il peut y avoir des différences de définition et de méthodologie entre l'enquête et la source extérieure qui risquent de compliquer l'interprétation de la différence observée. Autrement dit, la différence entre les estimations de l'enquête et celles de la source extérieure peut être due à d'autres causes que la non-réponse.

80. La mesure des différences entre répondants et non-répondants est coûteuse. En principe, à condition de disposer de suffisamment de ressources, on s'imagine parfois que l'on peut obtenir des réponses des cas de non-réponse. Cependant, les ressources sont rarement suffisantes pour que l'on tente d'obtenir de telles réponses. À la place, on peut envisager une seconde phase ou sélectionner un double échantillon parmi les non-répondants et consacrer toutes les ressources restantes sur le budget de l'enquête à recueillir des données auprès de ce sous-échantillon.

81. Statistiquement, il y a quelques écrits sur l'échantillonnage à deux phases pour les cas de non-réponse qui traitent d'un certain nombre d'aspects de sa conception (voir, par exemple, Cochran, 1977, sect. 13.6). Lorsqu'une réponse complète est obtenue de la deuxième phase de l'échantillonnage, on peut déterminer la fraction optimale d'échantillonnage de la seconde phase, compte tenu des contraintes financières, qui permet de limiter la variance d'échantillonnage d'une estimation en deux phases de la moyenne.

4. Réduction et compensation des cas de non-réponse unitaire dans les enquêtes sur les ménages

82. Dans beaucoup de cas, on réduit le nombre de cas de non-réponse unitaire par des méthodes ad hoc qui semblent des moyens judicieux de réduire les taux de non-réponse. Récemment, des théories d'ensemble fondées sur des principes sociologiques et psychologiques ont été formulées [voir Groves et Couper (1998)], qui

pourraient donner naissance à des méthodes de réduction des cas de non-réponse grâce à une meilleure compréhension de la façon dont se produisent ces cas dans les enquêtes sur les ménages. Il n'entre pas dans le cadre de ce chapitre de décrire ces cadres théoriques globaux. En revanche, nous décrivons plusieurs techniques qui sont ressorties d'études expérimentales comme des moyens efficaces de réduire les cas de non-réponse.

83. Les visites répétées, ou « nouvelles tentatives » sont courantes dans les enquêtes par échantillonnage. Les enquêteurs n'essaient pas seulement une fois mais plusieurs de contacter un ménage ou une personne à interroger, afin d'obtenir des réponses à toutes leurs questions. Le nombre de ces nouvelles tentatives, leur programmation et les techniques employées par l'enquêteur pour persuader les personnes réticentes ou difficiles à contacter sont autant de sujets de recherche sur le terrain. Cela dit, il n'y a pas de panacée pour répondre à ces problèmes. Les différences observées d'un pays à un autre en ce qui concerne les taux de réponse, l'acceptation des enquêtes par le public et la mobilité des populations rendent impossible l'élaboration d'une théorie unique au sujet de ces nouvelles tentatives. La réceptivité du public à différents thèmes des enquêtes ne permet pas d'établir des normes applicables à différents types d'enquêtes, même à l'intérieur d'un même pays. Cela dit, il est toujours préférable d'utiliser les meilleurs enquêteurs pour les cas difficiles de non-réponse.

84. Il n'y a aucune preuve empirique qu'une technique particulière, y compris les nouvelles tentatives, produise des taux de réponse élevés dans les enquêtes sur les ménages. Souvent, on a recours à une combinaison de techniques. Les enquêtes qui font suite à une notification sous la forme d'un appel téléphonique ou d'une lettre, d'une correspondance personnalisée, d'informations sur le parrainage de l'enquête et d'illustrations de la façon dont seront utilisées les données ont montré qu'elles obtenaient des taux de réponse plus élevés. Les incitations ne sont pas toujours bien reçues dans les pays en développement et les pays en transition et sont découragées dans beaucoup de pays. En revanche, elles sont de plus en plus courantes dans les enquêtes effectuées dans les pays développés [voir Kulka (1995) pour une revue des travaux de recherche sur les techniques employées].

85. On peut aussi améliorer les taux de réponse en examinant les techniques employées par les enquêteurs. La formation qui prépare l'enquêteur à adapter son approche aux réactions qu'il suscite de la part des répondants peut améliorer sensiblement ces taux. Les primes versées aux enquêteurs dont les résultats ont dépassé les objectifs de l'enquête ont aussi des effets bénéfiques sur les taux de réponse.

86. Dans toute enquête sur les ménages, les cas de non-réponse unitaire sont inévitables. Lors de la conception de l'enquête, on ajuste souvent la taille de l'échantillon en prévision de ces cas, et on calcule des pondérations compensatoires pour permettre d'ajuster les estimations et l'analyse.

87. L'ajustement de la taille de l'échantillon en prévision des cas de non-réponse nécessite l'estimation avant la collecte de données du taux de non-réponse attendu. Cette estimation est souvent ad hoc ou propre à une enquête donnée, et fondée sur les données d'enquêtes précédentes auprès de la population visée, sur le thème de l'enquête et sur d'autres facteurs. Pour une enquête transversale ponctuelle, l'estimation nécessite souvent que l'on suppose que ce qui s'est passé lors d'enquêtes précédentes se reproduira dans l'enquête à venir. Pour les enquêtes transversales répétées où la même population est sélectionnée à intervalles réguliers ou irréguliers, on dispose d'emblée des données à utiliser pour estimer les taux de réponse prévus. Pour les enquêtes par panel, où les unités de l'échantillon sont suivies dans le temps, l'estimation nécessite que l'on prévoie non seulement la non-réponse initiale à la première vague, mais aussi

la non-réponse par attrition, c'est-à-dire les cas où des sujets qui ont coopéré lors des phases précédentes ne peuvent plus être interrogés lors des phases suivantes (par refus, parce que l'on n'arrive plus à les joindre ou pour d'autres raisons).

88. L'ajustement augmente la taille de l'échantillon pour des raisons de précision, de manière à le doter d'un nombre suffisant d'unités pour qu'il produise le résultat désiré. Disons, par exemple, que l'on a besoin de 1 000 entrevues avec des ménages et que l'on prévoit un taux de non-réponse de 20 %. Pour obtenir les 1 000 entrevues voulues, il faut un échantillon de $1\,000/(1 - 0,2) = 1\,250$. Un échantillon final de cette taille, dans la mesure où le taux de réponse est exact, produira approximativement le nombre voulu d'entrevues complètes. On répartit les unités à interroger entre les enquêteurs, à qui l'on donne pour instructions de s'efforcer de recueillir le plus grand nombre de réponses possible. Aucune substitution n'est autorisée.

89. L'autre moyen de faire face à la non-réponse unitaire est la substitution. Cette approche laisse à l'enquêteur le soin de décider de contacter une unité, c'est-à-dire à un jugement subjectif et non pas à une sélection objective par probabilité qui permettrait de déterminer quelles unités de l'échantillon contacter. Il se peut que les méthodes par substitution produisent des échantillons de la taille exacte voulue, mais il a été amplement prouvé [voir, par exemple, Stephan et McCarthy (1958), qui analyse une procédure non fondée sur la probabilité, l'échantillonnage par quotas] que les méthodes de sélection par substitution débouchent sur des échantillons qui ne correspondent pas nécessairement à la composition de la population.

90. Des ajustements statistiques peuvent être appliqués aux données finales de l'enquête, de manière à compenser en partie la distorsion potentielle liée à la non-réponse. La forme la plus courante de compensation nécessite la définition de pondérations d'ajustement pour non-réponse.

91. Les pondérations d'ajustement pour non-réponse exigent que l'on ait les mêmes informations pour l'ensemble des répondants et des non-répondants. Comme on sait peu de choses des non-répondants, le type de variables dont on dispose pour ce type d'ajustement est limité dans la plupart des cas d'enquêtes sur les ménages. Le plus souvent, les informations primaires dont on dispose au sujet des ménages non répondants ont trait à leur localisation géographique, c'est-à-dire à l'endroit où ils vivent.

92. Par exemple, supposons qu'une enquête sur les ménages utilise une méthode d'échantillonnage géographique selon laquelle les zones d'énumération du recensement sont choisies à la première phase de la sélection. Durant la collecte de données, on ne recueille pas de données auprès de tous les ménages choisis pour l'enquête dans une zone d'énumération donnée. Une simple formule d'ajustement de pondération pour non-réponse attribuerait de plus fortes pondérations à tous les ménages d'une zone d'énumération qui répondent, afin de compenser les cas de non-réponse de ménages de cette zone. Si 90 % des ménages d'une zone d'énumération répondent, les pondérations données aux ménages qui répondent sont relevées selon un facteur de $1/0,9 = 1,11$. Si dans une autre zone, le taux de réponse est de 80 %, ce facteur devient $1/0,8 = 1,25$. On applique le même facteur d'accroissement aux pondérations données à tous les ménages de la zone d'énumération qui répondent. Les ménages qui ne répondent pas sont retirés de l'échantillon final, ce qui équivaut à leur affecter une pondération nulle.

93. Dans certains cas, les ajustements de pondération peuvent être définis par comparaison de données administratives aux données des ménages répondant à l'enquête. Par exemple, les données administratives ont pu être utilisées pour sélectionner l'échantillon. On peut alors attribuer aux répondants de l'échantillon des pondéra-

tions de telle sorte que la distribution de ces pondérations pour certaines variables essentielles corresponde à la distribution qui ressort des données administratives.

94. On peut également effectuer les ajustements pour non-réponse sur la base d'un modèle. Lorsque la classification des ménages sélectionnés pour une enquête se limite à distinguer ceux qui ont répondu de ceux qui n'ont pas répondu et que l'on dispose de données pour les uns comme pour les autres, la situation de répondant ou de non-répondant peut être régressée sur les variables disponibles. On peut alors utiliser des coefficients de régression logistique pour prédire la probabilité de réponse des ménages. On peut utiliser sensiblement de la même façon l'inverse des probabilités prédites pour calculer une pondération, appelée parfois pondération de la propension à répondre. Comme les pondérations calculées directement à partir des probabilités prédites tendent à être très variables, les probabilités prédites sont souvent groupées en catégories, et on affecte une seule pondération à chaque catégorie en utilisant l'inverse du point médian, ou la probabilité prédite moyenne, ou le taux de réponse pondéré de la catégorie comme pondération.

5. Non-réponse ponctuelle et imputation

95. La non-réponse ponctuelle a donné lieu récemment à d'importants travaux de recherche [voir, par exemple, l'étude récente de Groves *et al.* (2002)]. Sur ce type de non-réponse, on dispose d'une grande quantité de données sur chaque cas de non-réponse. Ces données permettent de mieux comprendre ce phénomène, et offrent la possibilité de le mesurer, de le réduire et de le compenser à l'aide de modèles statistiques plus complexes.

96. Par exemple, supposons que 90 % des répondants à une enquête sur les ménages portant sur la disponibilité de services et de soins de santé donnent des réponses à toutes les questions, mais que 10 % d'entre eux répondent à toutes les questions sauf une concernant leurs revenus salariaux du mois précédent. Les informations fournies par les 90 % des répondants qui ont donné des réponses complètes peuvent être utilisées pour élaborer des modèles statistiques destinés à faire mieux comprendre le rapport entre la santé et les soins de santé et les revenus salariaux. Ces modèles peuvent alors être utilisés pour concevoir des méthodes destinées à réduire le niveau de non-réponse aux questions sur les revenus salariaux, afin de compenser, voire de prédire les valeurs manquantes concernant ces revenus.

97. Le remplacement des valeurs manquantes sur un point est appelé imputation; on l'utilise depuis des décennies dans les enquêtes. Voir Kalton et Kasprzyk (1986) et Brick et Kalton (1996) pour un examen des procédures d'imputation utilisées dans les enquêtes sur les ménages ou dans d'autres enquêtes. L'imputation est une procédure utilisée dans les enquêtes pour compenser les valeurs manquantes. Il s'agit essentiellement de remplacer une valeur manquante sur un point par une valeur que l'on prédit sur la base des autres informations dont on dispose sur le sujet interrogé (ménage ou personne, par exemple) ou sur d'autres sujets interrogés dans le cadre de l'enquête.

98. L'imputation peut se faire, par exemple, au moyen d'un modèle de régression. Pour une variable Y d'une enquête, on peut proposer un modèle pour Y qui permette de « prédire » Y en utilisant un ensemble de p autres variables X_1, \dots, X_p de l'enquête. Un tel modèle est défini par l'équation :

$$Y_i = \beta_0 + \beta_1 X_{1i} + \dots + \beta_p X_{pi} + \varepsilon_i$$

Ce modèle est adapté à l'ensemble de sujets pour lesquels la variable Y de l'enquête et les variables « prédictes » X_1, \dots, X_p ne sont pas manquantes. Dès lors, on peut prédire la valeur de Y pour les cas manquants en utilisant les paramètres estimés obtenus par l'adaptation du modèle ci-dessus. La valeur prédite de la variable Y pour la i^e unité est donnée par :

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \dots + \hat{\beta}_p X_{pi}$$

99. Ce modèle de régression pour l'imputation est utilisé sous plusieurs formes. La prédiction par régression peut comprendre une valeur « résiduelle » prédite à ajouter à la valeur prédite. Une technique appelée imputation séquentielle applique une forme d'imputation par régression qui ajoute effectivement une valeur résiduelle « empruntée » à un autre cas au fichier de données avec des valeurs similaires de X_1, \dots, X_p au cas à imputer.

100. Des travaux récents dans le domaine de l'imputation ont également porté sur le problème découlant du fait que l'imputation introduit une variabilité supplémentaire dans les estimations qui utilisent les valeurs imputées. Cette variabilité peut être déterminée par des procédures d'estimation de la variance telles que l'estimation de la variance de type « *jackknife* », par des modèles pour le processus d'imputation ou par une procédure à imputations multiples où l'imputation est répétée plusieurs fois et la variabilité entre les valeurs imputées est incluse dans l'estimation de la variance.

101. Certaines techniques peuvent être utilisées pour réduire le niveau de non-réponse ponctuelle dans une enquête. Les enquêteurs peuvent apprendre à évaluer toute réponse non codifiable ou incomplète donnée à une question. Les concepteurs ajoutent en fait des questions de suivi discrètes afin d'insister sur certains points lorsque la réponse est « Je ne sais pas » ou « Je ne veux pas répondre à cette question ». Par exemple, les questions sur le revenu ont des taux de non-réponse plus élevés que certaines autres. Les enquêtes sur les revenus ajoutent parfois une série de questions sur certains éléments de revenu qui offrent des « fourchettes » à l'intérieur desquelles le revenu peut être indiqué. Si le répondant refuse de répondre ou ne connaît pas le montant du revenu, les questions peuvent être du type : Le revenu est-il de plus de XXX unités ? Est-il compris entre YYY et XXX unités ? etc. Ces questions permettent d'établir des fourchettes à l'intérieur desquelles se situe le revenu.

102. Les organismes qui effectuent des enquêtes sur les ménages devraient examiner régulièrement la fréquence des cas de non-réponse ponctuelle à travers les divers points de l'enquête pour mesurer l'importance du problème. Les taux de non-réponse ponctuelle sont rarement publiés, sauf pour certains points essentiels. Il est souvent laissé à l'utilisateur le soin de déterminer dans quelle mesure la non-réponse ponctuelle risque de lui poser un problème pour son analyse. La documentation sur l'enquête devrait inclure des taux de non-réponse pour les points essentiels et pour ceux pour lesquels le taux de non-réponse est élevé.

REMERCIEMENTS

L'auteur tient à remercier Kenneth Coleman, candidat à la maîtrise ès sciences au Programme d'étude de la méthodologie des enquêtes à l'Université du Michigan, de son aide précieuse pour l'examen des méthodes d'enquête employées en Amérique latine et en Amérique du Sud.

RÉFÉRENCES

- Brick, J. M., et G. Kalton (1996). Handling missing data in survey research. *Statistical Methods in Medical Research*, vol. 5, pp. 215-238.
- Cochran, W. G. (1977). *Sampling Techniques*. 3^e éd. New York, John Wiley and Sons.
- Groves, R. M. (1989). *Survey Errors and Survey Costs*. New York, John Wiley and Sons
- _____, et M. P. Couper (1998). *Non-response in Household Interview Surveys*. New York, John Wiley and Sons.
- Groves, R. M., et al. (2002). *Survey Non-response*. New York, John Wiley and Sons.
- Kalton, G., et D. Kasprzyk (1986). The treatment of missing survey data. *Survey Methodology*, vol. 12, pp. 1-16.
- Kish, L., et I. Hess (1950). On non-coverage of sample dwellings. *Journal of the American Statistical Association*, vol. 53, pp. 509-524.
- Kulka, R. (1995). The use of incentives to survey « hard-to-reach » respondents: a brief review of empirical research and current research practices. *Seminar on New Directions in Statistical Methodology*. Statistical Policy Working Paper, n° 23. Washington, D.C., U.S. Office of Management and Budget, pp. 256-299.
- Lessler, J., et W. Kalsbeek (1992). *Non-sampling Error in Surveys*. New York, John Wiley and Sons.
- Lepkowski, James M. (1988). The treatment of wave non-response in panel surveys. Dans *Panel Survey Design and Analysis*, D. Kasprzyk, G. Duncan et M. P. Singh, eds. New York, John Wiley and Sons
- Marks, E. S. (1978). The role of dual system estimation in census evaluation. In *Developments in Dual System Estimation of Population Size and Growth*, K. J. Krotki, ed. Edmonton, Alberta, University of Alberta Press.
- Seligson, M. A., et J. Jutkowitz (1994). *Guatemalan Values and the Prospects for Democratic Development*. Arlington, Virginie, Development Associates, Inc.

Chapitre IX

L'erreur de mesure dans les enquêtes sur les ménages : sources et évaluation

DANIEL KASPRZYK

Mathematica Policy Research
Washington, D.C., États-Unis d'Amérique

RÉSUMÉ

Le présent chapitre décrit les principales sources d'erreur de mesure dans les enquêtes par échantillonnage et les méthodes utilisées généralement pour chiffrer ces erreurs. Quatre sources d'erreur — le questionnaire, le mode de collecte des données, l'enquêteur et le répondant — sont examinées, et une description de la façon dont ces erreurs se produisent à travers ces sources est présentée. Les méthodes employées pour chiffrer ces erreurs, telles que les expériences randomisées, les études sur la recherche cognitive, les études à mesures répétées et les études par vérification des dossiers, sont décrites, et des exemples sont donnés pour illustrer leur application.

Termes clés : erreur de mesure, sources de l'erreur de mesure, méthodes employées pour chiffrer l'erreur de mesure.

A. Introduction

1. Les enquêtes sur les ménages utilisent diverses méthodes de collecte des données. Le processus de collecte de ces données repose sur l'hypothèse selon laquelle les caractéristiques et les concepts mesurés sont définis avec précision, peuvent être déterminés par un ensemble de procédures bien définies et ont des valeurs vraies indépendamment de l'enquête. L'erreur de mesure est donc la différence entre la valeur d'une caractéristique donnée par le répondant et la valeur vraie (mais inconnue) de cette caractéristique. Autrement dit, l'erreur de mesure est liée à l'observation de la variable à travers le processus de collecte de données de l'enquête et est de ce fait appelée parfois « erreur d'observation » (Groves, 1989).

2. Le présent chapitre est basé sur un chapitre consacré à l'erreur de mesure dans un document de travail préparé par une sous-commission de la mesure et de la notification de la qualité des données d'enquêtes de la Commission fédérale de la méthodologie statistique des États-Unis (2001). Par conséquent, un grand nombre des références et des exemples sont tirés de la recherche effectuée aux États-Unis d'Amérique et dans d'autres pays développés. Néanmoins, l'analyse s'applique à toutes les enquêtes, quel que soit l'endroit où elles se déroulent. Ce chapitre devrait donc être tout aussi utile pour les enquêtes effectuées dans les pays en développement et les pays en transition.

3. Il existe de nombreux écrits sur les erreurs de mesure dans les enquêtes par échantillonnage [voir Biemer *et al.* (1991) et Lyberg *et al.* (1997)] pour un examen des questions importantes qui se posent au sujet de ces erreurs. Une erreur de mesure peut donner lieu à la fois à des distorsions et à des erreurs sur les variables (variance) dans une estimation effectuée sur des essais répétés de l'enquête. La distorsion de la mesure ou distorsion de la réponse se présente comme une tendance ou une orientation systématique de la différence entre les réponses à une question et les valeurs vraies. Par exemple, les répondants tendent parfois à omettre d'indiquer un revenu tiré d'une deuxième ou d'une troisième source, ce qui se traduit par l'indication d'un revenu inférieur à leur revenu réel. Il y a variance si les valeurs indiquées sont différentes lorsque la question est posée plus d'une fois aux mêmes unités (ménages, personnes, enquêteurs et questionnaires) qui sont les sources d'erreurs. Une variance simple reflète une variation aléatoire dans la réponse d'un répondant à une question posée plusieurs fois (c'est-à-dire des réponses différentes à la même question posée plusieurs fois au répondant). Les effets variables que les enquêteurs peuvent avoir sur les réponses qu'ils reçoivent peuvent être source d'erreur sur une variable; c'est ce que l'on appelle la variance liée à l'enquêteur. Cette variance est une forme de variance sur réponse corrélée qui se produit parce que les erreurs de réponse sont corrélées pour les unités de l'échantillon interrogées par le même enquêteur.

4. Les études décrivent plusieurs approches générales de l'étude de l'erreur de mesure. L'une de ces approches compare les réponses à une enquête à des données potentiellement plus exactes émanant d'une autre source. Ces données peuvent se situer au niveau de l'unité de l'échantillon comme dans une « étude par vérification des dossiers ». À titre de simple exemple, si on demande leur âge aux répondants, leurs réponses peuvent être vérifiées sur les registres de l'état civil. Cela dit, il nous faut admettre que, même dans ce cas simple, on ne peut supposer à coup sûr que ces registres sont exempts d'erreurs. Néanmoins, l'un des moyens d'étudier l'erreur de mesure dans une enquête par sondage est de comparer les réponses reçues aux données émanant d'autres sources indépendantes valables. Un autre moyen d'évaluer l'erreur de mesure en utilisant des données émanant d'une autre source est d'effectuer l'analyse au niveau d'ensemble, c'est-à-dire de comparer les estimations fondées sur l'enquête aux estimations de la population provenant d'une autre source. Une deuxième autre approche consiste à effectuer des mesures répétées sur certaines des unités de l'échantillon. Il s'agit généralement d'un programme de secondes entrevues qui permet de comparer les réponses de la première entrevue à celles d'une seconde entrevue effectuée peu après la première. Une troisième approche consiste à sélectionner des sous-échantillons aléatoires à partir de l'échantillon complet de l'enquête et à leur administrer des traitements différents, tels que les soumettre à d'autres questionnaires ou à d'autres questions, ou encore à différents modes de collecte de données. Enfin, on peut également évaluer l'erreur de mesure par des moyens qualitatifs. Les méthodes employées comprennent les groupes de réflexion et les laboratoires sous contrôle, tels que les laboratoires de recherche cognitive.

5. Ce chapitre décrit les principales sources d'erreur de mesure trouvées dans les échantillons d'enquêtes et leur mesure. La mise en place de procédures pour chiffrer l'erreur de mesure est coûteuse et souvent difficile à réaliser. C'est pour cette raison, et parce que cela est jugé préférable, que les responsables d'enquêtes s'efforcent davantage de contrôler les sources d'erreur par une bonne planification et de bonnes pratiques d'exécution d'enquêtes. Ces pratiques consistent notamment à tester les matériels, questionnaires et procédures de l'enquête, à mettre au point et à tester des concepts d'enquête bien définis et opérationnellement réalisables, à accorder une attention particulière au traitement des questions de collecte des données auprès des

sous-groupes difficiles à atteindre, à appliquer des normes rigoureuses au recrutement d'un personnel de terrain qualifié et à élaborer et mettre en œuvre des programmes de formation intensive et des instructions précises et clairement énoncées à l'intention de ce personnel. La limitation des erreurs autres que les erreurs d'échantillonnage et, en particulier, des erreurs de mesure, nécessite en soi un examen approfondi. [Voir, par exemple, le rapport publié par les Nations Unies (1982), qui comprend une liste récapitulative pour la limitation de ces erreurs dans les enquêtes sur les ménages. Ce chapitre ne traite pas cette question mais s'efforce plutôt de décrire les principales sources d'erreurs de mesure dans les enquêtes par échantillonnage et les moyens généralement employés pour les chiffrer.]

6. À l'instar de Biemer *et al.* (1991), on examinera quatre sources d'erreur : le questionnaire, le mode de collecte des données, l'enquêteur et le répondant. Une part importante de ce chapitre est consacrée à décrire comment se produit l'erreur de mesure dans les enquêtes par échantillonnage à travers ces sources. Vient ensuite un examen de certaines approches employées pour chiffrer l'erreur de mesure. Ces approches comprennent les expériences randomisées, les études sur la recherche cognitive, les études à mesures répétées et les études par vérification des dossiers. Pour chiffrer l'erreur de mesure, il faut toujours prendre des mesures supplémentaires avant, pendant et après l'enquête. Souvent, ces études sont limitées du fait du temps et des ressources qu'il faut leur consacrer. Or, les études de l'erreur de mesure sont très utiles à la fois pour évaluer le niveau de l'erreur dans l'enquête en cours et pour indiquer ce qui pourrait être amélioré pour les enquêtes futures. Elles sont particulièrement intéressantes pour les programmes d'enquêtes répétées.

B. Les sources de l'erreur de mesure

7. Biemer *et al.* (1991) font état de quatre principales sources d'erreur de mesure :

- *Le questionnaire* : l'effet de la conception du questionnaire, sa présentation visuelle, les thèmes qu'il couvre et le libellé des questions.
- *La méthode de collecte des données* : l'effet de la façon dont le questionnaire parvient au répondant (par exemple, par courrier, remis en main propre ou présenté dans un journal). Les répondants peuvent répondre différemment aux questions en présence d'un enquêteur, s'ils répondent d'eux-mêmes ou s'ils le font par un journal.
- *L'enquêteur* : l'effet que l'enquêteur peut avoir sur les réponses aux questions. L'enquêteur peut introduire une erreur dans les réponses s'il ne lit pas les questions comme il le devrait, s'il donne une interprétation inappropriée d'une réponse mal formulée ou en ajoutant des informations qui risquent de prêter à confusion pour le répondant ou de l'égarer.
- *Le répondant* : le fait que les répondants, en raison d'expériences, de connaissances et d'attitudes différentes, risquent d'interpréter différemment le sens des questions.

8. Ces quatre sources sont d'une importance critique pour la conduite de l'enquête par sondage. Le questionnaire est le moyen employé pour demander certaines informations au répondant. La méthode de collecte des données a trait à la façon dont le questionnaire parvient au répondant (de façon impersonnelle ou en main propre). L'enquêteur est celui qui apporte le questionnaire lorsque celui-ci est remis en main propre, et le répondant est celui qui le reçoit. Chacun peut introduire une erreur dans le processus de mesure. La plupart des enquêtes considèrent ces sources séparément, à supposer qu'elles les prennent en compte. Toutefois, il peut y avoir interaction entre

ces sources; par exemple, l'enquêteur et le répondant peuvent, par leur comportement l'un envers l'autre, introduire des erreurs qui n'apparaissent pas si l'on considère ces sources séparément. On examinera ci-après les diverses façons selon lesquelles des erreurs de mesure peuvent être introduites par ces quatre sources.

1. Les effets du questionnaire

9. Le questionnaire est l'instrument qu'utilise l'enquêteur pour recueillir des informations auprès d'un répondant. Au cours des 20 dernières années, les principes fondamentaux de conception des questionnaires, jugés naguère comme relevant plus de l'art que de la science, ont été le thème d'abondants ouvrages (Sirken *et al.*, 1999; Schwarz, 1997; Sudman, Bradburn et Schwarz, 1996; Bradburn et Sudman, 1991). Le questionnaire ou ses caractéristiques, c'est-à-dire la façon dont les questions sont libellées ou la façon dont le questionnaire lui-même est conçu, peuvent influencer sur la façon dont un individu répond aux questions. Dans la présente section, nous examinerons les différents moyens par lesquels le questionnaire peut introduire des erreurs dans le processus de collecte de données.

Les problèmes de spécifications

10. Lors de la planification d'une enquête, il se pose souvent des problèmes dus au fait que les objectifs de la recherche et les informations que le questionnaire est censé recueillir sont ambigus, mal définis ou incohérents. La façon dont les questions sont formulées ne permet pas de recueillir les informations nécessaires à la réalisation des objectifs de la recherche. Il peut se poser des problèmes de spécification des données parce que le questionnaire est mal conçu et les instructions données sont mal rédigées, parce que les définitions sont ambiguës ou parce que le concept étudié est difficile à mesurer. Par exemple, le questionnaire peut chercher à savoir quels ont été « les soins prénatals dispensés pendant la grossesse » mais sans préciser de quelle grossesse ni de quelle période il s'agit. L'ambiguïté peut caractériser certaines questions aussi élémentaires que celle-ci : Combien d'emplois occupez-vous ? si l'on ne précise pas la nature de l'emploi — temporaire ou permanent et à plein temps ou à temps partiel. Les concepts analytiques composites tels que le revenu total d'une personne, ne sont pas nécessairement révélés dans leur totalité si les divers éléments qui composent le revenu ne sont pas identifiés et définis pour le répondant.

Le libellé des questions

11. Pour que le répondant les interprète correctement, il est important que les questions soient libellées clairement et avec précision. Comme le questionnaire est une forme de communication entre l'enquêteur et le répondant, il y a de nombreuses sources d'erreur potentielles. Premièrement, il se peut que le concepteur du questionnaire n'ait pas clairement défini le concept qu'il/elle cherche à mesurer. Ensuite, même si le concept est clairement formulé, il peut ne pas être convenablement représenté par la question ou l'ensemble de questions dont il fait l'objet, et même s'il est clairement et fidèlement représenté, l'interprétation qu'en tire le répondant peut très bien ne pas être celle recherchée. Les différences de langue et de culture ou les différences d'expérience et de contexte entre le concepteur du questionnaire et le répondant peuvent être cause d'une mauvaise interprétation des questions. Ces différences peuvent se révéler particulièrement importantes dans les pays en développement et les pays en transition qui comptent plusieurs groupes ethniques différents. Vaessen *et al.* (1987) examinent les problèmes linguistiques qui se posent lors des enquêtes effectuées dans des pays à plusieurs langues.

12. Il y a au moins deux niveaux dans la compréhension d'une question posée dans une enquête par sondage. Le premier niveau est celui de la simple compréhension du sens littéral de la question. Le répondant comprend-il les mots utilisés dans le libellé de la question ? Peut-il fournir des informations qui correspondent à sa compréhension de ces mots et offrir une réponse significative ? Cela dit, pour répondre à une question, le répondant doit aussi supputer l'intention qui motive le questionnaire; autrement dit, pour répondre à la question, il doit être capable d'en déterminer le sens pragmatique (Schwarz, Groves et Schuman, 1995). C'est ce second élément qui rend le libellé des questions plus difficile et plus complexe qu'une simple formulation nécessitant un faible niveau de compréhension de la question à lire. Pour produire un instrument bien conçu, il faut la contribution des répondants, c'est-à-dire leur interprétation et leur compréhension des questions. Les méthodes de recherche cognitive offrent souvent un moyen utile d'obtenir cette contribution (voir section C.2).

Longueur des questions

13. Le bon sens et un bon style d'écriture militent en faveur de questions courtes et simples, de nature à faciliter leur interprétation. Cependant, la recherche a montré que des questions plus longues peuvent susciter des réponses plus détaillées des répondants, du moins pour ce qui est du comportement à l'égard de certains symptômes et des visites chez le médecin (Marquis et Cannell, 1971) et de la consommation d'alcool et de drogue (Bradburn et associés, 1979). Les questions plus longues fournissent plus d'informations et de détails de nature à aider le répondant à se rappeler et plus de temps pour réfléchir aux informations demandées.

Longueur du questionnaire

14. Les chercheurs et les analystes veulent toujours poser le plus de questions possible, tandis que le concepteur de l'enquête se rend compte que des erreurs peuvent être introduites dans un questionnaire trop long. Le répondant peut perdre sa concentration ou se fatiguer selon son état (âge ou maladie, par exemple), la clarté du sujet, son rapport avec l'enquêteur, la conception du questionnaire et la méthode employée pour l'entrevue.

Ordre des questions

15. Les chercheurs ont remarqué que l'ordre dans lequel sont posées les questions influe sur la réponse (Schuman et Presser, 1981), en particulier dans les enquêtes sur les opinions et attitudes. On a constaté des phénomènes d'assimilation, lorsque les réponses suivantes vont dans le même sens que les précédentes, et de contraste, lorsque les réponses suivantes vont dans le sens opposé à celui des précédentes. Les répondants peuvent aussi utiliser les informations découlant de questions précédentes pour l'interprétation des mêmes termes dans leurs réponses aux questions suivantes.

Catégories de réponse

16. Les catégories de questions peuvent influencer sur les réponses en donnant une impression au répondant de ce que le concepteur du questionnaire juge important. Le répondant peut s'imaginer que les catégories figurant sous une rubrique sont celles que le concepteur juge les plus importantes. Cela peut prêter à confusion quant à l'intention recherchée par la question si le répondant en retire l'impression que les catégories de réponses ne semblent pas appropriées. L'ordre des catégories peut également influencer sur les réponses. Les répondants peuvent choisir la facilité lors d'une

entrevue et répondre systématiquement par la même note sur une échelle de réponses, choisir les notes qui se présentent en premier ou inversement en dernier.

17. L'effet produit par l'ordre des catégories de réponses peut aussi être fonction de la façon dont le questionnaire est rempli. Si le répondant remplit seul le questionnaire, les catégories de réponses figurant en premier dans la liste ont plus de chances d'être retenues et approuvées (effet de primauté), car le répondant a plus de temps pour y répondre. Si les questions sont posées par un enquêteur, les catégories figurant par la suite ont plus de chances d'être retenues (effet de récence).

Formats ouverts et fermés

18. Un format de question qui donne aux répondants la possibilité de choisir entre un ensemble spécifié de réponses (format fermé) peut produire des réponses différentes de celles où le répondant n'a pas ce choix (format ouvert) [Bishop *et al.*, 1988]. Une réponse donnée a plus de chances d'apparaître dans un format ouvert que lorsqu'elle figure comme option dans un format fermé (Bradburn, 1983). Le format fermé peut rappeler aux répondants d'inclure une information qu'ils risqueraient autrement d'oublier. Les choix de réponses peuvent indiquer aux répondants le niveau ou le type de réponses jugées appropriées [voir, par exemple, Schwarz, Groves et Schuman (1995) et Schwarz et Hippler (1991)].

Format du questionnaire

19. La « physionomie » du questionnaire rempli sans la présence de l'enquêteur, c'est-à-dire le format et la présentation du questionnaire, peut favoriser l'exactitude des réponses ou, au contraire, y faire obstacle. Le fait que les répondants peuvent être troublés par un questionnaire mal conçu peut se traduire par des malentendus ou contribuer à une interprétation erronée des questions et des instructions. Jenkins et Dillman (1997) proposent des principes pour la conception des questionnaires à remplir seul pour la population des États-Unis. Il faut faire preuve de prudence dans la transposition de ces principes à d'autres pays sans tenir compte des facteurs culturels et linguistiques propres à ce pays.

2. Les effets liés au mode de collecte des données

20. Le choix du mode de collecte des données est une décision qui implique une variété de questions. Les ressources financières jouent souvent un rôle important; néanmoins, la teneur du questionnaire, la population cible, les taux de réponse visés, la durée de la période de collecte des données et l'erreur de mesure prévue sont autant de facteurs importants dans le processus de décision du mode de collecte à choisir. Si les progrès de la technologie ont engendré un accroissement de l'usage du téléphone pour la collecte des données, les autres options sont nombreuses pour la façon d'effectuer l'enquête. Lyberg et Kasprzyk (1991) présentent un aperçu des différentes méthodes de collecte ainsi que des sources d'erreur de mesure liées à chacune de ces méthodes. On trouvera ci-dessous un résumé de cet aperçu.

Entrevue face à face

21. Les entrevues face à face sont la principale méthode de collecte de données dans les pays en développement et les pays en transition. Le plus souvent, un enquêteur présente un questionnaire structuré aux répondants et inscrit leurs réponses sur le questionnaire. L'utilisation de cette méthode d'entrevue personnelle avec papier et crayon [*paper and pencil personal interview (PAPI)*] a un long passé. Grâce aux progrès

récents dans la production d'ordinateurs portables légers, les entrevues face à face sont à présent consignées sur ordinateur [*computer-assisted personal interviewing* (CAPI)]. Les enquêteurs visitent les répondants chez eux et effectuent l'entrevue en consignnant les réponses sur l'ordinateur portable plutôt que sur un questionnaire imprimé. Voir Couper *et al.* (1998) pour un examen des questions qui se posent au sujet de l'entrevue assistée par ordinateur « CAPI ». L'avantage le plus évident de la méthode CAPI a trait au contrôle de la qualité et à la réduction des erreurs de réponse. Les enquêteurs entrent les réponses sur un fichier informatique. Le logiciel de l'entrevue garantit que les questions à sauter sont bien omises et que les réponses sont inscrites et éditées de manière à assurer leur validité au moment de l'entrevue; cela permet d'économiser du temps et des ressources au stade du dépouillement des données de l'enquête.

22. Les entrevues face à face permettent d'effectuer des entrevues complexes, d'utiliser des aides visuelles pour aider le répondant à répondre aux questions, d'établir un rapport entre les enquêteurs qualifiés et les répondants et d'assurer l'obtention de réponses plus complètes et plus exactes. Toutefois, les enquêteurs peuvent parfois influencer sur les réponses des répondants et créer ainsi une distorsion dans les estimations de l'enquête ou un effet de variance lié à l'enquêteur (voir section C.3). Les enquêteurs peuvent infléchir les réponses par une combinaison de traits liés à leur personnalité et à leur comportement. On peut craindre certains traits ou comportements indésirables. Les répondants peuvent hésiter à signaler ces traits ou comportements à l'enquêteur. DeMaio (1984) note que le facteur d'acceptabilité sociale semble comporter deux éléments : l'idée que certaines choses sont bonnes et d'autres mauvaises, et le fait que les répondants veulent faire « bonne impression » et tendent à donner des réponses dans ce sens.

23. L'autre source possible d'erreur de mesure liée à l'entrevue face à face dans les enquêtes sur les ménages a trait à la présence possible d'autres membres de la famille au moment de l'entrevue. Ces membres peuvent influencer sur les réponses du répondant, surtout si les questions sont jugées délicates. Par exemple, il peut être difficile à un répondant de répondre à des questions concernant l'utilisation de drogues illicites en présence d'un autre membre de la famille. Même des questions à première vue innocentes peuvent être perçues comme délicates en présence d'un autre membre de la famille (par exemple, les questions liées aux antécédents conjugaux ou à la fécondité, les questions portant sur le passé posées en présence du conjoint).

Questionnaires remplis sans la présence d'un enquêteur

24. Les sources d'erreur de mesure lorsque les questionnaires sont remplis en l'absence d'un enquêteur sont différentes de celles des entrevues face à face. Les questionnaires remplis par le répondant sans l'aide d'un enquêteur présentent les caractéristiques suivantes : bien entendu, il n'y a aucune influence d'un enquêteur, et moins de risques d'effets « d'acceptabilité sociale ». Ces questionnaires permettent également de poser des questions sur des sujets délicats ou potentiellement dangereux sans embarrasser le répondant. Ils offrent aussi l'avantage de pouvoir, s'il y a lieu, être administrés simultanément à plus d'un répondant au sein de la famille (Dillman, 1983). En revanche, ils peuvent souffrir d'une distorsion systématique si la population cible est composée d'individus avec peu ou pas d'instruction, ou d'individus qui ont du mal à lire et à écrire. On peut observer cette distorsion dans les réponses aux questions « ouvertes » qui risquent d'être moins précises et détaillées que celles obtenues dans des enquêtes effectuées par des enquêteurs. Cette méthode de collecte de données n'est pas toujours idéale dans les pays où le taux d'alphabétisme est faible; toutefois, même si la population cible a un niveau d'instruction raisonnablement élevé, il peut arriver

que les répondants commettent des erreurs d'interprétation des questions ou des instructions. Généralement, les taux de réponse ponctuelle sont plus faibles avec les questionnaires remplis sans l'aide d'un enquêteur, mais lorsque des réponses sont données aux questions, les données sont de meilleure qualité. Les questionnaires remplis sans l'aide d'un enquêteur, peut-être plus que les autres modes de collecte de données, gagnent à ce que le questionnaire soit bien conçu et à ce que les questions y soient présentées clairement et intelligiblement. L'un des types particuliers de ce questionnaire est celui où les répondants doivent remplir le questionnaire et le retourner par courrier (Dillman, 1978; 1991; 2000).

Enquêtes par journal

25. Les enquêtes par journal sont des formulaires remplis par le répondant lui-même, qui sont utilisés pour des sujets qui nécessitent un compte rendu détaillé du comportement sur une période donnée (par exemple, les dépenses, l'emploi du temps et l'usage fait de la télévision). Pour limiter ou éviter les défaillances de la mémoire, le répondant est encouragé à utiliser un journal et à y inscrire les réponses au sujet d'un événement peu après qu'il se produit. Pour que cette formule réussisse, il faut que le répondant s'emploie activement à consigner les informations et à remplir un formulaire généralement fastidieux. Il faut également que la population cible sache lire et interpréter les questions figurant sur le journal, condition qui écarte les pays à faible taux d'alphabétisme. La procédure de collecte de données nécessite généralement qu'un enquêteur contacte le répondant pour que celui-ci remette le journal, pour obtenir sa coopération et pour lui expliquer les procédures d'enregistrement des données. L'enquêteur revient après un certain délai pour récupérer le journal et, s'il n'a pas été rempli, pour aider le répondant à le remplir.

26. Lyberg et Kasprzyk (1991) identifient un certain nombre de sources d'erreur de mesure pour ce mode de collecte de données. Par exemple, les répondants qui ne se soucient guère ou pas du tout d'enregistrer les événements risquent fort de ne pas les entrer dans le journal quand ils sont encore frais dans leur mémoire. Le journal lui-même, par sa conception et sa forme et par la complexité des questions, peut poser au répondant un certain nombre de problèmes pratiques. En outre, les répondants peuvent modifier leur comportement par suite de l'utilisation du journal; par exemple, le fait d'avoir à énumérer ses achats dans un journal de dépenses peut l'inciter à modifier son comportement d'acheteur. On trouvera des analyses des erreurs de mesure dans les enquêtes sur les dépenses et, en particulier, dans le journal tenu pour les besoins de ces enquêtes, dans Neter (1970) et Kantorowitz (1992). Silberstein et Scott (1991) offrent des comparaisons des données obtenues par les entrevues face à face et par les enquêtes par journal.

Observation directe

27. En tant que méthode de collecte de données, l'observation directe nécessite que l'enquêteur collecte les données en utilisant ses sens (vue, ouïe, toucher, goût) ou des moyens concrets de mesure. Cette méthode est utilisée dans beaucoup de domaines, par exemple, dans les enquêtes sur l'agriculture, pour estimer les rendements des récoltes (estimation visuelle) et dans les enquêtes sur les ménages, pour évaluer la qualité des logements des répondants. Les observateurs introduisent des erreurs de mesure de façons similaires à celles par lesquelles les erreurs sont introduites par les enquêteurs; par exemple, il peut arriver que les observateurs ne comprennent pas certains concepts et commettent des erreurs de perception des informations à enre-

gistrer, ou modifient leur mode d'enregistrement des informations dans le temps par tendance à la facilité ou par fatigue.

3. Les effets liés à l'enquêteur

28. L'enquêteur joue un rôle critique dans beaucoup d'enquêtes par échantillonnage. En tant qu'élément fondamental du processus de collecte de données, sa performance peut influencer sur la qualité des données de l'enquête. Cela dit, l'enquêteur n'est que l'une des composantes du processus de collecte, dont le chercheur ou le responsable de l'enquête peut chercher à contrôler la performance; c'est pourquoi des stratégies sont apparues — par sélection et recrutement, formation et suivi du comportement professionnel — pour limiter l'erreur liée à son rôle (Fowler, 1991). En raison de différences individuelles, chaque enquêteur aborde l'enquête d'une façon différente; par exemple, les enquêteurs ne posent pas nécessairement les questions exactement comme elles sont libellées, ne suivent pas toujours correctement des instructions concernant les questions à sauter ou ne justifient pas toujours les réponses de la manière appropriée. Ils peuvent ne pas suivre les instructions à la lettre, soit à dessein soit parce que ces instructions ne sont pas claires. Inconsciemment, les enquêteurs peuvent infléchir le ton de leur voix, ou afficher des changements d'attitude.

29. Chaque enquêteur peut introduire des erreurs par excès ou par défaut. Lorsque ces erreurs se compensent plus ou moins à travers toutes les entrevues, il peut se produire une légère distorsion imputable à l'enquêteur. Toutefois, les erreurs commises par tel ou tel enquêteur peuvent être grossières et aller dans le même sens, ce qui crée de grosses distorsions imputables à ces enquêteurs. Les variations des distorsions entre les différents enquêteurs donnent lieu à ce que l'on appelle la variance liée à l'enquêteur, qui peut avoir de sérieuses répercussions sur la précision des estimations de l'enquête.

Variance liée à l'enquêteur

30. Au début des années 1960, Kish (1962) a élaboré une approche utilisant un coefficient de corrélation inhérent à l'enquêteur, qu'il a appelé ρ , pour évaluer l'effet de la variance liée à l'enquêteur sur les estimations de l'enquête. La grandeur ρ , qui est définie comme le ratio de la composante variance liée à l'enquêteur à la variance totale d'une variable de l'enquête, est estimée par une simple analyse de la variance.

31. Dans les enquêtes effectuées face à face, ρ est généralement d'environ 0,02 pour la plupart des variables. Bien que cette valeur de ρ soit faible, son effet sur la précision de l'estimation peut être sensible. La variance de la moyenne de l'échantillon est multipliée par $1 + \rho (n - 1)$, où n est l'ampleur de la charge de travail moyenne de l'enquêteur. Pour une valeur de ρ de 0,02 et une charge de travail de 10 entrevues, la variance augmente de 18 %, et une charge de travail de 25 entrevues produit une variance de 48 % plus élevée. Autrement dit, même de faibles valeurs de ρ peuvent réduire sensiblement la précision des statistiques de l'enquête. Pour des raisons économiques, les enquêteurs ont généralement des charges de travail très lourdes. De ce fait, un enquêteur qui produit une distorsion systématique influe sur les résultats obtenus auprès d'un nombre considérable de répondants et peut avoir un effet sensible sur la variance.

Caractéristiques de l'enquêteur

32. Les travaux de recherche n'aident guère à définir les caractéristiques d'un bon enquêteur. Au Royaume-Uni de Grande-Bretagne et d'Irlande du Nord, Collins

(1980) n'a trouvé aucun fondement à recommander que le recrutement d'enquêteurs se concentre plutôt sur les femmes que sur les hommes, ou parmi les personnes de classe moyenne, ou encore parmi les personnes d'âge moyen plutôt que sur les jeunes ou les personnes âgées. Weiss (1968), étudiant un échantillon de mères vivant de l'aide sociale à New York, a confirmé l'exactitude de plusieurs points et constaté que la similarité entre l'enquêteur et le répondant en ce qui concerne l'âge, l'éducation et la condition socio-économique ne produit pas de meilleurs résultats. Sudman *et al.* (1977) ont étudié les attentes de l'enquêteur sur la difficulté d'obtention d'informations délicates et constaté certaines faiblesses dans le rapport entre les difficultés prévues et effectivement rencontrées lors de l'entrevue. Groves (1989) a examiné un certain nombre d'études et conclu, d'une façon générale, qu'il peut se produire des effets démographiques lorsque les mesures sont liées aux caractéristiques démographiques, mais pas dans le cas contraire; par exemple, si les questions ont trait à la race, cela peut avoir un effet selon la race de l'enquêteur.

Méthodes de contrôle des erreurs commises par l'enquêteur

33. Dans une certaine mesure, le responsable de l'enquête peut contrôler les erreurs commises par l'enquêteur en veillant à ce qu'une formation soit dispensée à celui-ci, en le supervisant et en modifiant sa charge de travail. Un programme de formation suffisamment long pour couvrir les aptitudes et les techniques à employer lors des entrevues et pour fournir des informations sur certaines enquêtes données permet d'assurer un certain degré de normalisation du processus d'entrevue (Fowler, 1991). Nombreux sont ceux qui pensent qu'en normalisant les procédures d'entrevue, on réduit les effets liés à l'enquêteur.

34. La supervision et le suivi de l'action de l'enquêteur, par l'observation et la tenue de statistiques de performance et l'identification des questions qui sont sources de problèmes, est une autre composante du système de contrôle de qualité du travail de l'enquêteur. Des programmes de seconde entrevue et des observations sur le terrain sont organisés pour évaluer la prestation de chaque enquêteur. Des observations sont effectuées sur le terrain au moyen de longues listes de codification ou de guides détaillés à l'usage des observateurs qui permettent au superviseur de vérifier si les procédures sont convenablement suivies. Par exemple, l'observation peut porter sur l'apparence et le comportement de l'enquêteur, la façon dont il se présente ou dont il présente l'enquête, la façon dont il pose les questions et recueille les réponses, l'utilisation de cartes et de moyens de contrôle neutres et la bonne utilisation du manuel de l'enquêteur. Dans d'autres cas, on peut enregistrer des bandes magnétiques (audio ou vidéo) et coder et analyser le comportement de l'enquêteur (Lyberg et Kasprzyk, 1991).

35. Un autre moyen de réduire l'effet de la variance liée à l'enquêteur est d'abaisser sa charge de travail moyenne; toutefois, cela implique que l'on puisse utiliser des enquêteurs supplémentaires de même niveau de compétence. Groves et Magilavy (1986) étudient la charge de travail optimale des enquêteurs en regard des coûts de recrutement et de formation des enquêteurs, des coûts des entrevues et de l'ampleur de l'influence personnelle de l'enquêteur. Comme cette influence varie selon les statistiques, il est très difficile de déterminer avec certitude ce qui constitue une charge de travail optimale.

36. On peut réduire les effets liés aux enquêteurs en évitant les problèmes découlant de la conception des questionnaires, en leur donnant des instructions et des définitions claires et sans ambiguïté, en leur apprenant à suivre les instructions et en limitant le recours aux aptitudes personnelles (et variables) de l'enquêteur à obtenir des réponses aux questions posées.

4. Les effets liés au répondant

37. Les répondants peuvent contribuer à introduire des erreurs de mesure en ne donnant pas des réponses exactes. Groves (1989) examine à la fois l'influence des modèles traditionnels d'entrevue (Kahn et Cannell, 1957) et des sciences de la connaissance sur les réponses à l'enquête. Hastie et Carlston (1980) identifient cinq stades séquentiels dans la formation et la communication de réponses par les répondants :

- *Codification de l'information*, c'est-à-dire le processus de mémorisation ou de conservation de la connaissance.
- *Compréhension des questions posées*, ce qui implique la connaissance des mots et expressions figurant dans le questionnaire ainsi que la compréhension par le répondant du but de l'enquête, du contexte et du libellé des questions, ainsi que du comportement de l'enquêteur qui les lui pose.
- *Extraction de l'information enfouie dans la mémoire*, c'est-à-dire l'effort que fait le répondant pour retrouver les informations pertinentes dans sa mémoire.
- *Choix de la réponse à donner à une question*, entre plusieurs réponses possibles, sur la base des informations retrouvées dans la mémoire.
- *Communication de la réponse*, ce qui implique une présentation exacte de l'information pertinente que le répondant a extrait de sa mémoire et l'aptitude du répondant à bien articuler sa réponse.

38. De nombreux aspects du processus d'enquête influent sur la qualité des réponses qui ressortent de ce processus en cinq stades. Voici quelques exemples des facteurs qui conditionnent les effets liés au répondant :

Règles applicables aux répondants

39. Les règles qui déterminent les critères d'admissibilité utilisés pour identifier la (les) personne(s) désignée(s) pour répondre au questionnaire jouent un rôle important dans le processus de réponse. S'il s'agit de recueillir des informations sur les ménages, la connaissance des réponses aux questions peut varier selon les divers répondants potentiels du ménage. Les enquêtes pour la collecte d'informations sur certaines personnes au sein des ménages sélectionnés peuvent utiliser les réponses obtenues directement auprès des intéressés ou des réponses obtenues auprès d'autres répondants. Les différences entre ces deux options varient selon le sujet (par exemple, l'obtention de réponses directement auprès des intéressés sont préférables dans les enquêtes sur le comportement). Une étude des Nations Unies (1982) décrit le résultat d'un test pilote des effets des réponses obtenues auprès d'autres répondants sur les questions démographiques dans l'Enquête démographique effectuée en Turquie. Blair, Menon et Bickart (1991) présentent une bibliographie des travaux de recherche sur la comparaison entre l'obtention de réponses directement auprès des intéressés et auprès d'autres répondants.

Questions

40. Le libellé et la complexité de la question et la conception du questionnaire peuvent influencer sur la compréhension de la question par le répondant (voir section B.1 pour plus de détails). L'inclination du répondant à donner des réponses exactes varie selon les types de questions posées, selon la difficulté à déterminer la réponse à donner et selon la perception qu'a le répondant de l'opportunité de répondre.

Enquêteurs

41. L'apparence de l'enquêteur (par exemple, son âge, son sexe, le ton et l'inflexion de sa voix et le rythme de son parler) peuvent aussi influencer sur la compréhension de la question.

Mémoire

42. Généralement, le temps émousse l'aptitude à se souvenir de faits ou d'événements. La mémoire s'estompe, et les répondants ont plus de mal à se rappeler un événement survenu longtemps auparavant. Ainsi, dans certains pays étudiés dans le cadre de l'Enquête mondiale sur la fécondité, les naissances récentes sont souvent datées avec plus de précision que celles qui remontent à une période plus éloignée (Singh, 1987). Les concepteurs de l'enquête ont intérêt à sonder des périodes qui tendent à réduire l'erreur carrée moyenne totale du point de vue de l'erreur d'échantillonnage et d'éventuelles distorsions; par exemple, Huang (1993) a constaté que la précision augmentait avec la taille de l'échantillon et qu'en prolongeant la période de référence de quatre mois à six mois, on ne compensait pas l'accroissement de la distorsion découlant de la perte de mémoire. Eisenhower, Mathiowetz et Morganstein (1991) examinent la question de l'utilisation de moyens d'aider à se souvenir (par exemple, l'utilisation de calendriers, de cartes, de journaux) pour réduire les distorsions dues à la perte de mémoire. Mathiowetz (2000) rapporte les résultats d'une méta-analyse testant l'hypothèse selon laquelle la qualité des rapports rétrospectifs est fonction de la durée écoulée depuis l'événement relaté.

Télescopage

43. Il y a télescopage lorsque les répondants rendent compte d'un événement comme étant survenu pendant la période de référence alors qu'il s'est produit en dehors de cette période. Des moyens mnémotechniques (par exemple, la conduite d'une entrevue initiale pour fixer une date de référence, ou l'utilisation d'une date importante ou d'un événement important comme point de départ de la période de référence) peuvent être employés pour réduire les risques de télescopage (Neter et Waksberg, 1964).

Enquêtes par panel ou longitudinales

44. D'autres facteurs liés aux répondants contribuent à engendrer des erreurs dans les enquêtes par panel ou enquêtes longitudinales. Par exemple, des mesures trompeuses peuvent apparaître lorsqu'un répondant donne des réponses différentes à la même question ou à des questions similaires et que ces réponses ne sont pas le reflet d'un changement réel. Kalton, Kasprzyk et McMillen (1986) offrent des exemples d'erreurs de mesure dans les vagues successives d'une enquête longitudinale. Ils citent l'âge, la race, le sexe, le secteur d'activité et la profession comme variables où l'erreur de mesure a été observée aux États-Unis dans l'Enquête sur les revenus et la participation à un programme. Le profil de qualité de cette enquête examine cette erreur et d'autres erreurs de mesure identifiées dans cette enquête (États-Unis, Bureau of the Census, 1998). Les techniques de « *dependent interviewing* », où les réponses à l'entrevue précédente sont utilisées dans l'entrevue actuelle, permettent de réduire l'incidence des changements trompeurs. Hill (1994) a constaté que le « *dependent interviewing* » permettait une nette amélioration des mesures du changement de profession et de secteur d'activité, mais qu'il pouvait aussi manquer de signaler un véritable changement, de sorte qu'il convient de l'utiliser avec discernement. Mathiowetz et McGonagle (2000) examinent les pratiques actuelles de l'entrevue assistée par ordina-

teur ainsi que les manifestations empiriques de l'impact du « *dependent interviewing* » sur la qualité des données.

45. Le conditionnement du panel ou distorsion liée aux dates d'entrevue avec l'échantillon est une autre source d'erreur potentielle dans les enquêtes par panel. Le conditionnement se réfère aux variations de la réponse d'un répondant interrogé plusieurs fois. Woltman et Bushery (1977) ont étudié la distorsion liée aux dates d'entrevue avec l'échantillon dans l'Enquête nationale auprès des victimes de crimes aux États-Unis, comparant les rapports de victimes ayant divers degrés d'expérience des panels (c'est-à-dire ayant fait l'objet d'un certain nombre d'entrevues précédentes) interrogées durant le même mois. Ils ont constaté d'une façon générale une diminution du nombre de crimes signalés à mesure qu'augmentait le nombre d'entrevues précédentes. Kalton, Kasprzyk et McMillen (1989) ont également examiné cette source d'erreur.

C. Approches de l'évaluation quantitative de l'erreur de mesure

46. Il y a plusieurs approches générales de l'évaluation quantitative de l'erreur de mesure. Pour étudier les distorsions de la mesure, divers traitements, tels que l'utilisation d'autres questionnaires ou d'autres questions ou d'un mode différent de collecte des données, peuvent être appliqués à des sous-échantillons de l'échantillon de l'enquête choisis au hasard. L'erreur de mesure peut être étudiée d'un point de vue quantitatif, dans le cadre de groupes de réflexion ou de laboratoires de recherche cognitive. Une autre approche consiste à effectuer des mesures répétées sur la même unité de l'échantillon, telles que celles qui sont effectuées dans un programme d'entrevues répétées. Enfin, il y a les « études des archives », qui permettent de comparer les réponses données lors de l'enquête à des données plus précises émanant d'une autre source d'estimation de l'erreur de mesure. Ces approches sont examinées ci-après.

1. Expériences effectuées au hasard

47. L'expérience effectuée au hasard est souvent utilisée comme méthode d'estimation des erreurs de mesure. Les chercheurs appellent cette méthode d'une variété de noms tels qu'échantillons interpénétrés, expériences sur échantillons fractionnés, expériences sur panels fractionnés, expériences sur semi-échantillons aléatoires et expériences sur bulletins fractionnés. Différents traitements liés à l'erreur spécifique mesurée sont administrés à des semi-échantillons aléatoires de conception identique. Pour l'étude d'erreurs variables, différentes entités considérées comme pouvant être la source de l'erreur sont incluses et comparées (par exemple, plusieurs enquêteurs différents pour des estimations de la variance liée à l'enquêteur). Pour l'étude des distorsions, on ne compare généralement que deux ou trois traitements (par exemple, deux modes différents de collecte des données), l'une des méthodes ayant la préférence. Les essais sur le terrain, effectués avant l'enquête, comprennent souvent des expériences aléatoires destinées à évaluer diverses méthodes et procédures et divers questionnaires possibles.

48. Par exemple, une expérience effectuée au hasard peut être utilisée pour tester l'effet de la longueur du questionnaire. Des unités d'échantillonnage sont affectées au hasard à un groupe d'un ensemble de deux groupes dont l'un reçoit une version « brève » des questions et l'autre une version « longue ». À supposer que l'on dispose d'une source de données indépendante, on peut alors comparer les réponses de chaque groupe aux estimations de la source de données, qui sont censées être exactes et

fiables. De même, on peut évaluer les effets de l'ordre dans lequel sont posées les questions en inversant cet ordre dans un autre questionnaire utilisé sur des échantillons aléatoires. Cette méthode a été utilisée pour une enquête en République dominicaine, effectuée dans le cadre du Programme d'enquêtes démographiques et sanitaires; le questionnaire de base a été utilisé sur les deux tiers de l'échantillon et le questionnaire expérimental sur le troisième tiers de l'échantillon. Il s'agissait de déterminer les différences entre les réponses obtenues aux deux ensembles de questions (Westoff, Goldman et Moreno, 1990).

2. Méthodes de recherche cognitive

49. Au cours des 20 dernières années, l'utilisation des méthodes de recherche cognitive pour réduire l'erreur de mesure s'est rapidement développée. Ces méthodes ont d'abord été utilisées pour analyser les processus de pensée des répondants, mais on les emploie de plus en plus comme supplément aux essais traditionnels sur le terrain (Schwarz et Sudman, 1996; Sudman, Bradburn et Schwarz, 1996). Les répondants fournissent des informations au concepteur du questionnaire sur la façon dont ils interprètent les questions. Cette approche demande l'intervention de beaucoup de personnes et s'avère d'un coût élevé par répondant; c'est pourquoi l'analyse cognitive est effectuée sur de petits échantillons. Le concepteur du questionnaire doit tenir compte du fait que ses conclusions révèlent des problèmes potentiels sans être pour autant nécessairement représentatifs des éventuels répondants à l'enquête.

50. Les méthodes les plus couramment utilisées s'appuient sur des protocoles verbaux (Willis, Royston et Bercini, 1991). Il est demandé aux répondants de répondre au projet de questionnaire et d'indiquer comment ils interprètent chaque question. Un enquêteur examine certains mots, expressions, la façon de sauter certaines questions et autres éléments particuliers du questionnaire sur lesquels il cherche à obtenir la réaction du répondant. Il demande aux répondants de lui indiquer tout ce qui ne leur paraît pas clair. Il peut leur demander de le faire en même temps qu'ils répondent au questionnaire (en réfléchissant en même temps « à haute voix ») ou au cours d'une séance organisée après coup (réflexion rétrospective « à haute voix »). Le concepteur peut ajouter des vérifications afin d'évaluer la clarté de différents points ou éléments du questionnaire lors d'entrevues ultérieures. Cette technique a pour avantage de ne pas être sujette à des distorsions imposées par l'enquêteur, mais elle a pour inconvénient de ne pas donner de bons résultats avec les répondants qui n'ont pas l'habitude d'exprimer leur pensée et éprouvent de la gêne à le faire.

51. Selon une technique apparentée, l'enquêteur interroge le répondant sur un certain aspect du questionnaire immédiatement après que celui-ci a répondu à un ensemble de questions (Nolin et Chandler, 1996). Cette approche est moins sujette à l'état d'esprit du répondant et à son aptitude à exprimer ses pensées, mais elle limite la recherche aux parties du questionnaire sur lesquelles le concepteur estime qu'il peut poser des questions. Elle peut aussi introduire une distorsion liée à l'enquêteur car c'est lui qui décide des vérifications. Dans la mesure où l'approche suivie pour la vérification diffère de la conduite d'une entrevue, certains la jugent artificielle (Willis, 1994).

52. D'autres approches permettent au répondant de remplir le questionnaire sans qu'il y ait lieu de faire appel à des groupes de réflexion. Les groupes de réflexion offrent l'avantage de l'interaction entre les membres du groupe, ce qui peut favoriser l'exploration de domaines qui pourraient être entièrement laissés de côté dans des entrevues en tête à tête.

53. La convocation de petits groupes d'experts réunis pour critiquer le questionnaire peut être un moyen efficace de cerner les problèmes que pose celui-ci (Czaja et Blair, 1996). Des spécialistes de la conception de questionnaires d'enquête et/ou des thèmes traités par ceux-ci reçoivent le questionnaire plusieurs jours avant une réunion avec ses concepteurs. Au cours d'une réunion de groupe, chacun examine et commente le questionnaire en prenant chaque question une par une.

54. Les méthodes de recherche cognitive sont aujourd'hui largement utilisées pour la conception de questionnaires et pour réduire l'erreur de mesure dans les enquêtes effectuées dans les pays développés. Sudman, Bradburn et Schwarz (1996) résumant leurs principales conclusions concernant la méthodologie de l'enquête. Tucker (1997) examine les questions méthodologiques qui se posent dans l'application de la psychologie cognitive à la recherche sur les enquêtes.

3. Études fondées sur une seconde entrevue

55. Au cours de la seconde entrevue — entrevue répétée d'une enquête par sondage — les mêmes questions (ou un sous-ensemble de celles-ci) sont posées à la même unité. Ces secondes entrevues sont effectuées à une ou plusieurs des fins suivantes :

- Pour identifier les enquêteurs qui falsifient les données;
- Pour identifier les enquêteurs qui interprètent mal les procédures et ont besoin d'une formation supplémentaire;
- Pour estimer la variance sur une réponse simple;
- Pour estimer une distorsion sur une réponse.

56. Les deux premières fins nous informent au sujet des erreurs de mesure qui sont le fait de l'enquêteur. Les deux dernières nous éclairent sur les erreurs de mesure qui sont dues à la fois aux quatre sources (à savoir l'enquêteur, le questionnaire, le répondant et la méthode de collecte des données).

57. Pour chacun de ces quatre types de seconde entrevue, on examine les conditions particulières à remplir lors de la conception [voir Forsman and Schreiner (1991)]. Par ailleurs, on examinera également certaines méthodes d'analyse des données résultant de la seconde entrevue, ainsi que les limitations quant aux résultats.

Seconde entrevue effectuée pour identifier les enquêteurs qui falsifient les données

58. Les enquêteurs peuvent falsifier de plusieurs façons les résultats de l'enquête; par exemple, l'enquêteur peut fabriquer des réponses à certaines ou à l'ensemble des questions, ou refuser délibérément de suivre les procédures de l'enquête. Pour détecter les cas de falsification, on sélectionne un échantillon pour une seconde entrevue qui est généralement effectuée par un membre du personnel d'encadrement. On peut calculer le taux de falsification, qui se définit comme le pourcentage d'enquêteurs qui falsifient les réponses tel qu'il ressort de ces secondes entrevues. D'après Schreiner, Pennie et Newbrough (1988), ce taux était de 0,4 % lors de l'Enquête récente sur la population aux États-Unis, de 0,4 % lors de l'Enquête nationale sur la criminalité aux États-Unis et de 6,5 % lors de l'enquête sur l'habitat et les logements vacants à New York; ces trois enquêtes ont été effectuées par le Bureau du recensement des États-Unis.

Seconde entrevue effectuée pour évaluer les enquêteurs

59. Certains programmes de secondes entrevues sont conçus pour identifier les enquêteurs qui ne fournissent pas un niveau de prestation acceptable. Il s'agit d'identifier les enquêteurs qui ne comprennent pas les procédures de l'enquête, afin de leur dispenser une formation supplémentaire. La plupart des aspects de ce type de seconde entrevue sont identiques à ceux de la seconde entrevue effectuée pour identifier les enquêteurs qui falsifient les données. Des tables de tolérance, fondées sur la théorie du contrôle de la qualité statistique, peuvent être utilisées pour déterminer si le nombre de différences observées lors de la seconde entrevue après comparaison dépasse la limite acceptable. Au Bureau du recensement des États-Unis, les programmes de seconde entrevue utilisent des niveaux de tolérance acceptables variant entre 6 et 10 % (Forsman et Schreiner, 1991).

Seconde entrevue effectuée pour évaluer la variance simple des réponses

60. La seconde entrevue effectuée pour évaluer la variance simple des réponses est une répétition des procédures de l'entrevue initiale. Toutes les directives et procédures de l'entrevue initiale sont reprises dans toute la mesure du possible lors de la seconde entrevue. L'échantillon utilisé pour cette seconde entrevue est un sous-échantillon représentatif de l'échantillon original. On utilise les enquêteurs, la méthode de collecte des données, les règles applicables aux répondants et les questionnaires de l'entrevue initiale. Dans la pratique, ces règles ne sont pas toujours suivies; par exemple, si le questionnaire original est trop long, on utilise un sous-ensemble de questions de ce questionnaire. On n'essaie *pas* d'aplanir les différences entre la première et la seconde entrevue.

61. L'une des statistiques estimées à partir de la variance simple des réponses est le taux de différence brut [*gross difference rate (GDR)*], qui est le carré de la différence moyenne entre les réponses à la première et à la seconde entrevue. Le *GDR* divisé par deux est une estimation sans distorsion de la variance simple des réponses [*simple response variance (SRV)*]. Pour les caractéristiques qui offrent deux choix possibles, le *GDR* est égal au pourcentage de cas qui ont présenté des réponses différentes entre la première et la seconde entrevue. Brick, Rizzo et Wernimont (1997) offrent des règles générales pour l'interprétation de la variance des réponses mesurée par le *GDR*.

62. Une autre statistique est l'indice d'incohérence [*index of inconsistency (IOI)*], qui mesure la proportion de la variance de la population totale attribuée à la simple variance des réponses. Autrement dit :

$$IOI = \frac{GDR}{s_1^2 + s_2^2}$$

où s_1^2 est la variance de l'échantillon pour l'entrevue initiale et s_2^2 la variance de l'échantillon pour la seconde entrevue.

63. On interprète souvent la valeur de l'*IOI* comme suit :

- Un *IOI* de moins de 20 % est une variance relative *faible* des réponses;
- Un *IOI* compris entre 20 et 50 % est une variance relative *modérée* des réponses;
- Un *IOI* de plus de 50 % est une variance relative *élevée* des réponses.

64. Les mesures de la variance des réponses, le *GDR* et l'*IOI*, donnent des informations aux utilisateurs des données sur la fiabilité et la cohérence des réponses aux questions d'une enquête. On trouvera des exemples de l'utilisation du *GDR* et de

l'IOI pour certaines variables d'une enquête sur la fécondité au Pérou dans l'étude des Nations Unies (1982) sur l'erreur autre que l'erreur d'échantillonnage dans les enquêtes sur les ménages. Dans le cadre de la seconde phase du programme d'Enquêtes démographiques et sanitaires, un programme de secondes entrevues destiné à évaluer la cohérence des réponses à l'échelon national a été organisé au Pakistan sur un sous-échantillon de femmes interrogées lors de l'enquête principale (Curtis et Arnold, 1994). Westoff, Goldman et Moreno (1990) décrivent une étude des secondes entrevues effectuée dans le cadre du programme d'Enquêtes démographiques et sanitaires en République dominicaine, qui mérite d'être notée en raison de la nécessité d'accepter plusieurs compromis, notamment l'obligation de restreindre les secondes entrevues à quelques zones géographiques et à un sous-ensemble de la population cible. Des enquêtes effectuées en Inde avec seconde entrevue, afin de mesurer la variance des réponses, sont décrites par l'United States Bureau of the Census (1985), qui examine les procédures d'évaluation des recensements.

65. Feindt, Schreiner et Bushery (1997) décrivent les efforts d'une enquête périodique en vue d'améliorer continuellement les questionnaires en utilisant un programme de secondes entrevues. Lorsque les réponses présentent des taux élevés d'incohérence révélés lors des secondes entrevues, on peut entreprendre des travaux de recherche en vue d'améliorer le questionnaire en utilisant les méthodes de recherche cognitive. Ces méthodes permettent d'identifier les causes de problèmes et de suggérer d'éventuelles solutions. Lors du cycle d'entrevues suivant, une seconde entrevue peut être effectuée à l'aide des questions révisées afin de déterminer si la fiabilité s'est améliorée. Ce processus est ensuite répété pour les autres questions à problème.

Seconde entrevue effectuée pour évaluer la distorsion des réponses

66. La seconde entrevue effectuée pour mesurer la distorsion des réponses a pour objet de déterminer quelles sont les réponses vraies ou exactes sur un sous-échantillon représentatif de l'échantillon initial. Pour obtenir les vraies réponses, on utilise les enquêteurs et les superviseurs les plus expérimentés. De plus, pour la seconde entrevue, on peut soit utiliser le répondant qui a le plus de connaissances au sein du ménage soit interroger les membres de la famille individuellement. On utilise les questions de l'entrevue initiale, et l'on compare les deux réponses en présence du répondant pour déterminer « la vérité ». On peut aussi utiliser une série de questions tests à la place des questions originales, afin d'obtenir des réponses exactes, puis comparer les réponses en présence du répondant. Pour une analyse des enquêtes avec seconde entrevue effectuées afin d'obtenir des estimations de la distorsion des réponses, voir le rapport décrivant les procédures d'évaluation du recensement publié par l'United States Bureau of the Census (1985).

67. La mise en concordance pour déterminer la vérité a ses limites. Les répondants peuvent délibérément donner de fausses informations tant lors de la première entrevue que lors de la seconde, de sorte que la comparaison ne donne toujours pas une estimation « vraie ». Dans une étude de la qualité effectuée sur les données de la seconde entrevue de l'Enquête récente sur la population aux États-Unis, Biemer et Forsman (1992) ont constaté que jusqu'à 50 % des erreurs que contenait l'entrevue initiale n'avaient pas été détectées lors de la seconde entrevue.

68. On estime la distorsion des réponses en calculant le taux de différence net [*net difference rate (NDR)*], c'est-à-dire la différence moyenne entre la réponse à l'en-

trevue initiale et celle de la seconde entrevue, censée représenter la réponse « vraie ». En l'occurrence :

$$NDR = \frac{1}{n} \sum_{i=1}^n (y_{O_i} - y_{T_i})$$

où n est la taille de l'échantillon soumis à la seconde entrevue; y_o est la réponse à l'entrevue initiale; et y_T est la réponse à la seconde entrevue après comparaison, censée être la réponse « vraie ».

69. Le *NDR* indique le taux de précision d'une réponse à un questionnaire et permet également d'identifier les questions qui donnent des résultats faussés. Il faut tenir compte de cette distorsion lorsque l'on analyse les données et que l'on interprète les résultats. Brick *et al.* (1996) ont utilisé la seconde entrevue de façon intensive pour mieux comprendre le point de vue du répondant et les raisons qui motivent ses réponses et justifient une estimation de leur distorsion. Bien qu'ils n'aient utilisé qu'un petit échantillon, les auteurs ont conclu que cette méthode permettait de détecter les distorsions et de les mesurer.

4. Études consacrées à la vérification des résultats

70. L'étude de la vérification des résultats compare les réponses données à un questionnaire par certains éléments de l'échantillon aux valeurs obtenues d'une source extérieure, dont on suppose généralement qu'elle contient les valeurs vraies des variables de l'enquête. Ces études permettent d'estimer la distorsion des réponses résultant de l'effet combiné des quatre sources d'erreur de mesure (l'enquêteur, le questionnaire, le répondant et la méthode de collecte des données).

71. Groves (1989) décrit les trois types d'études de la vérification des résultats :

- La contre-vérification des données;
- La vérification directe des données;
- La vérification complète des données.

72. Dans l'étude de contre-vérification, l'échantillon de l'enquête est choisi à partir d'une source de données exactes sur les caractéristiques de l'étude. L'estimation de la distorsion de la réponse est alors fondée sur une comparaison des réponses à l'enquête aux données de cette source.

73. Souvent, la source de données est une liste d'unités (de ménages ou de personnes) présentant une caractéristique particulière — par exemple, les unités qui reçoivent une certaine forme d'aide de l'État. Dans ce cas, l'étude de contre-vérification des données ne mesure pas les erreurs dues aux unités qui affirment posséder une caractéristique qu'elles ne possèdent pas. Ces études ne peuvent mesurer que la proportion des unités de l'échantillon qui indiquent à juste titre qu'elles possèdent cette caractéristique et la proportion de celles qui l'indiquent de façon erronée. Par exemple, une étude de contre-vérification des données a été effectuée par l'United States Law Enforcement Assistance Administration (1972) pour mesurer les erreurs dans les cas de victimisation. Les archives de la police ont fait l'objet d'un échantillonnage et des victimes indiquées sur des dossiers ont été contactées. Lors des entrevues, les victimes ont signalé 74 % des crimes enregistrés dans les archives de la police.

74. Dans l'étude par vérification directe, des archives extérieures contenant des informations exactes sur les répondants sont étudiées une fois que les réponses à l'enquête ont été recueillies. Les estimations de la distorsion des réponses sont fondées sur

une comparaison des réponses données à l'enquête aux valeurs consignées dans les archives. Les études par vérification directe permettent de mesurer les cas où une caractéristique est indiquée de façon erronée. L'une des difficultés de ce type d'études est qu'elles nécessitent que l'on contacte les organismes qui conservent ces archives et que l'on obtienne des répondants l'autorisation de vérifier leurs informations. Si la réponse au questionnaire de l'enquête indique que l'unité ne présente pas une caractéristique donnée, il peut être difficile de fouiller les archives à la recherche de cette unité. Par conséquent, les études par vérification directe sont limitées dans leur aptitude à mesurer les omissions. Chaney (1994) décrit une étude par vérification directe qui compare les indications données par des enseignants de leurs qualifications aux archives universitaires. Cette étude a révélé que les informations fournies par les enseignants eux-mêmes sur leurs diplômes et les années où ils les ont obtenus ainsi que leurs principaux domaines d'études étaient exacts pour la plupart; en revanche, l'indication des cours suivis et des crédits obtenus comportait plus d'inexactitudes.

75. L'étude par vérification complète des données combine des aspects de la contre-vérification et de la vérification directe. Un échantillon est sélectionné à partir d'un cadre couvrant l'ensemble de la population et les données de toutes origines sur les cas de l'échantillon sont localisées. Ainsi on peut mesurer les erreurs dues à des omissions ou à des données erronées en comparant les réponses données par les répondants à toutes les archives (c'est-à-dire aux données émanant du cadre d'échantillonnage ainsi que de sources extérieures). Bien que ce type de vérification des données permette d'éviter les faiblesses de la contre-vérification et de la vérification directe, il n'exige pas une base de données couvrant toutes les unités de la population et tous les événements correspondants pour ces unités. Marquis et Moore (1990) donnent une description détaillée de la conception et de l'analyse d'une étude de vérification complète effectuée pour estimer les erreurs de mesure de l'Enquête sur les revenus et la participation aux programmes aux États-Unis. Dans cette étude, les données de l'enquête sur les montants des prestations perçues au titre de huit programmes de prestations fédérales et des États dans quatre États ont été comparées aux données des archives administratives de ces programmes. Le Profil de qualité des enquêtes (United States Bureau of the Census, 1998) présente un résumé de cette conception et de cette analyse.

76. Ces trois types d'études de la vérification des données ont leurs limites liées aux trois hypothèses suivantes qui, dans la pratique, sont irréalistes et ne sont jamais justifiées : premièrement, que les archives sont exemptes d'erreurs de couverture, de non-réponse ou de données manquantes; deuxièmement, que leurs dossiers individuels sont complets, exacts et exempts d'erreurs de mesure; et troisièmement, qu'il n'est pas commis d'erreurs lorsque l'on compare les réponses à l'enquête aux données sur les répondants.

77. On peut estimer la distorsion d'une réponse pour une caractéristique donnée en prenant la différence moyenne entre la réponse au questionnaire et la donnée correspondante des archives, selon la formule :

$$\text{Distorsion d'une réponse} = \frac{1}{n} \sum_{i=1}^n (Y_i - X_i)$$

où : n est la taille de l'échantillon de l'étude de vérification; Y_i = la réponse à l'enquête pour la i^{e} personne de l'échantillon; et X_i = la grandeur donnée par les archives pour la i^{e} personne de l'échantillon.

78. Les mesures de la distorsion des réponses d'une étude de vérification par comparaison aux archives fournissent des informations sur la précision d'une question

et permettent de repérer les questions qui donnent des estimations faussées. Ces mesures peuvent aussi être utilisées pour évaluer diverses options concernant certaines caractéristiques de l'enquête, telles que la conception du questionnaire, les périodes de réminiscence, les méthodes de collecte des données et les techniques de « *bounding* ». Par exemple, Cash et Moss (1972) donnent les résultats d'une étude de contre-vérification dans trois comtés de la Caroline du Nord en ce qui concerne la notification d'accidents de la circulation. Des entrevues ont été organisées dans des ménages dont certains membres avaient été identifiés comme ayant été impliqués dans un accident de la circulation dans les douze mois qui avaient précédé l'entrevue. L'étude a montré que, alors que 3,4 % seulement des accidents survenus au cours des trois mois qui avaient précédé l'entrevue n'avaient pas été notifiés, cette proportion montait à plus de 27 % pour ceux qui s'étaient produits de 9 à 12 mois avant l'entrevue.

5. Études de la variance liée à l'enquêteur

79. Pour étudier la variance liée à l'enquêteur, les affectations des enquêteurs doivent se faire au hasard, de manière que les résultats obtenus par différents enquêteurs puissent être associés à ces derniers.

80. On estime la variance liée à l'enquêteur en affectant chacun de ceux-ci à des répondants différents mais similaires, c'est-à-dire ayant les mêmes attributs vis-à-vis des variables de l'enquête. Dans la pratique, on assure cette équivalence par distribution aléatoire. L'échantillon est divisé en sous-ensembles constitués de façon aléatoire, dont chacun est représentatif de la même population, et chaque enquêteur travaille sur l'un de ces sous-ensembles. Ainsi, à son niveau, il effectue une enquête sur un sous-échantillon présentant essentiellement les mêmes attributs que l'échantillon entier, mais à une échelle plus restreinte. O'Muircheartaigh (1982) décrit la méthode employée pour l'Enquête mondiale sur la fécondité pour mesurer la variance de la réponse imputable aux enquêteurs et donne des estimations de cette variance pour les enquêtes effectuées au Lesotho (1984a) et au Pérou (1984b).

81. Pour les entrevues face à face, les affectations des enquêteurs sont conçues géographiquement de manière à éviter des coûts de déplacement trop lourds. Les zones d'affectation sont d'une taille à la mesure de la charge de travail de chaque enquêteur et sont identifiées et attribuées par paires à deux enquêteurs. À l'intérieur de chacune de ces zones, une moitié des unités d'habitation sélectionnées est attribuée au hasard à chacun d'entre eux. Ainsi, chaque enquêteur effectue des entrevues dans deux zones d'affectation et chacune de ces zones est répartie entre deux enquêteurs. L'expérience (la comparaison des résultats obtenus par les deux enquêteurs dans chacune des deux zones qui leur sont affectées) est répétée autant de fois qu'il y a d'équipes de deux enquêteurs. Bailey, Moore et Bailar (1978) présentent un exemple de l'interpénétration pour les entrevues en tête à tête dans l'Enquête nationale auprès des victimes de crimes dans huit villes des États-Unis.

6. Évaluation codée du comportement

82. On peut évaluer la performance des enquêteurs tant lors de leur formation que lors des enquêtes en procédant à une évaluation codée de leur comportement. Des observateurs qualifiés observent un échantillon d'enquêteurs; les aspects à coder des entrevues ou les entrevues elles-mêmes sont enregistrés sur bande magnétique et la codification se fait au moyen de ces bandes. Des codes sont attribués à chacune des principales activités verbales de l'enquêteur et à certains comportements tels que la façon dont il pose les questions et dont il les exprime, et enregistre les réponses. Par

exemple, des codes peuvent servir à évaluer la façon dont l'enquêteur lit les questions, à déterminer si les questions sont posées correctement et dans leur intégralité, si elles subissent de légères modifications ou omissions et si l'enquêteur pose la question dans ses propres termes ou de façon incomplète. Le système de codification permet d'indiquer si le libellé employé a pu orienter le répondant vers une réponse particulière, a ajouté des précisions à la question ou a laissé au répondant le choix de sa réponse, si les réponses ont été résumées de façon précise ou imprécise, et si divers autres comportements de l'enquêteur étaient ou non appropriés. Les résultats codifiés indiquent dans quelle mesure l'enquêteur a employé les méthodes qui lui avaient été enseignées; autrement dit, son comportement est jugé « incorrect » ou « inapproprié » lorsqu'il lui a été recommandé de l'éviter. Pour établir et maintenir un haut niveau de fiabilité de cette codification de la part de chaque enquêteur, un second codeur doit coder séparément un sous-échantillon d'entrevues.

83. Le système de codification du comportement permet d'indiquer aux nouveaux enquêteurs quelles sont les techniques d'entrevue qui sont acceptables et quelles sont celles qui ne le sont pas, et peut servir de base à l'examen du travail de terrain par les enquêteurs et les superviseurs et à l'analyse des problèmes révélés par la codification. En outre, il permet d'évaluer la performance de l'enquêteur, qui peut être comparée à celles d'autres enquêteurs et à sa propre performance lors d'autres entrevues codifiées (Cannell, Lawson et Hauser, 1975).

84. Oksenberg, Cannell et Blixt (1996) décrivent une étude où le comportement des enquêteurs a été enregistré, codifié et analysé aux fins d'identification des problèmes imputables respectivement à l'enquêteur et au répondant lors de l'Enquête nationale sur les frais médicaux effectuée par l'United States Agency for Health Care Research and Quality. Cette étude devait déterminer si le comportement de l'enquêteur s'était écarté des principes et des techniques qui lui avaient été inculqués lors de sa formation. Les auteurs ont signalé que les enquêteurs s'étaient souvent écartés du libellé des questions et avaient souvent posé celles-ci d'une manière qui tendait à orienter la réponse. Les enquêteurs n'avaient pas suffisamment aidé les répondants, et lorsqu'ils l'avaient fait, leur aide avait eu tendance à être orientée ou inappropriée.

D. Conclusions : l'erreur de mesure

85. L'erreur de mesure se produit durant le processus de collecte des données. Quatre principales sources d'erreur ont été identifiées comme inhérentes à ce processus : le questionnaire, la méthode ou le mode de collecte des données, l'enquêteur et le répondant. Pour déterminer l'existence et l'ampleur d'une erreur de mesure donnée, il faut une planification préalable et un examen minutieux. À moins qu'il s'agisse d'enquêtes de faible ampleur (c'est-à-dire effectuées sur un échantillon limité), il faut des études spéciales qui exigent une distribution aléatoire des sous-échantillons, des secondes entrevues et des vérifications des résultats. Ces études sont généralement coûteuses et nécessitent la présence d'un statisticien pour l'analyse des données. Néanmoins, si l'on craint à juste titre que la question ne puisse être convenablement résolue lors de la préparation de l'enquête ou si la source d'erreur est particulièrement conséquente, il importe que les responsables de l'enquête prennent des dispositions pour concevoir des études spéciales afin de chiffrer la principale et la plus problématique source d'erreur.

86. On ne saurait trop insister sur l'importance des études pour comprendre et chiffrer l'erreur de mesure dans une enquête. Cela est particulièrement critique si les concepts de l'enquête sont nouveaux et complexes. La qualité des analyses qu'effectuent les utilisateurs dépend de la qualité des données recueillies et de la compréhens-

sion de la nature et des limites de ces données. Les études de l'erreur de mesure exigent le plus grand sérieux de la part de l'enquête, car elles sont coûteuses et laborieuses. Toutefois, ce sérieux ne se termine pas avec la mise en œuvre et la conduite de ces études. Celles-ci doivent être analysées et leurs résultats consignés de façon que les analystes puissent effectuer leur propre évaluation de l'effet de l'erreur de mesure sur les résultats. On peut trouver des études axées sur l'analyse de tests et d'expérience et sur l'évaluation de la qualité des données dans différents rapports techniques et méthodologiques [voir, par exemple, les rapports méthodologiques et analytiques produits par le programme d'Enquêtes démographiques et sanitaires (Stanton, Abderrahim et Hill, 1997; Institute for Resource Development/Macro Systems Inc., 1990; Curtis, 1995)]. Enfin, les résultats des études de l'erreur de mesure sont un important moyen d'améliorer l'exécution d'enquêtes à venir. L'amélioration sensible des mesures passe pour beaucoup par la connaissance des résultats d'enquêtes précédentes. L'amélioration de la qualité des données des enquêtes à venir nécessite l'engagement des spécialistes de la recherche dans ce domaine.

RÉFÉRENCES

- Bailey, L., T. F. Moore et B. A. Bailer (1978). An interviewer variance study for the eight impact cities of the National Crime Survey Cities Sample. *Journal of the American Statistical Association*, vol. 73, pp. 16-23.
- Biemer, P. P., et G. Forsman (1992). On the quality of reinterview data with application to the current population survey. *Journal of the American Statistical Association*, vol. 87, pp. 915-923.
- Biemer, P. P., *et al.*, eds. (1991). *Measurement Errors in Surveys*. New York, John Wiley and Sons.
- Bishop, G. F., *et al.* (1988). A comparison of response effects in self-administered and telephone surveys. dans *Telephone Survey Methodology*, R. M. Groves *et al.*, eds. New York, John Wiley and Sons, pp. 321-340.
- Blair, J., G. Menon et B. Bickart (1991). Measurement effects in self vs. proxy responses to survey questions: an information-processing perspective. Dans *Measurement Errors in Surveys*, P. Biemer *et al.*, eds. New York, John Wiley and Sons, pp. 145-166.
- Bradburn, N. M. (1983). Response Effects. Dans *Handbook of Survey Research*, P. H. Rossi, J. D. Wright et A. B. Anderson, eds. New York, Academic Press, pp. 289-328.
- _____, et S. Sudman (1991). The current status of questionnaire design. Dans *Measurement Errors in Surveys*, P. Biemer *et al.*, eds. New York, John Wiley and Sons, pp. 29-40.
- _____, et associés (1979). *Improving Interviewing Methods and Questionnaire Design: Response Effects to Threatening Questions in Survey Research*. San Francisco, Californie, Jossey-Bass.
- Brick, J. M., L. Rizzo et J. Wernimont (1997). *Reinterview Results for the School Safety and Discipline and School Readiness Components*. Washington, D.C., United States Department of Education, National Center for Education Statistics. NCES 97-339.
- Brick, J. M., *et al.* (1996). *Estimation of Response Bias in the NHES: 95 Adult Education Survey*. Working Paper, No. 96-13. Washington, D.C., United States Department of Education, National Center for Education Statistics.
- Cannell, C. F., S. A. Lawson et D. L. Hauser (1975). *A Technique for Evaluating Interviewer Performance*. Ann Arbor, Michigan, University of Michigan, Survey Research Center.

- Cash, W. S., et A. J. Moss (1972). Optimum recall period for reporting persons injured in motor vehicle accidents. *Vital and Health Statistics*, vol. 2, n° 50. Washington, D.C., Public Health Service.
- Chaney, B. (1994). *The Accuracy of Teachers' Self-reports on Their Post Secondary Education: Teacher Transcript Study, Schools and Staffing Survey*. Working Paper, No. 94-04. Washington, D.C., United States Department of Education, National Center for Education Statistics.
- Collins, M. (1980). Interviewer variability: a review of the problem. *Journal of the Market Research Society*, vol. 22, n° 2, pp. 77-95.
- Couper, M. P., et al., eds. (1998). *Computer Assisted Survey Information Collection*. New York, John Wiley and Sons.
- Curtis, S. L. (1995). *Assessment of the Data Quality of Data Used for Direct Estimation of Infant and Child Mortality in DHS-II Surveys*. Occasional Papers, No. 3. Calverton, Maryland, Macro International, Inc.
- _____, et F. Arnold (1994). *An Evaluation of the Pakistan DHS Survey Based on the Reinterview Survey*. Occasional Papers No. 1. Calverton, Maryland, Macro International, Inc.
- Czaja R., et J. Blair (1996). *Designing Surveys: A Guide to Decisions and Procedures*. Thousand Oaks, Californie, Pine Forge Press (a Sage Publications company).
- DeMaio, T. J. (1984). Social desirability and survey measurement: a review. Dans *Surveying Subjective Phenomena*, C. F. Turner et E. Martin, eds. New York, Russell Sage, pp. 257-282.
- Dillman, D. A. (1978). *Mail and Telephone Surveys: The Total Design Method*. New York, John Wiley and Sons.
- _____. (1983). Mail and other self-administered questionnaires. Dans *Handbook of Survey Research*, P. Rossi, R. A. Wright et B. A. Anderson, eds. New York, Academic Press, pp. 359-377.
- _____. (1991). The design and administration of mail surveys. *Annual Review of Sociology*, vol. 17, pp. 225-249.
- _____. (2000). *Mail and Internet Surveys: The Tailored Design Method*. New York, John Wiley and Sons.
- Eisenhower, D., N. A. Mathiowetz et D. Morganstein (1991). Recall error: sources and bias reduction techniques. Dans *Measurement Errors in Surveys*, P. Biemer et al., eds. New York, John Wiley and Sons, pp. 127-144.
- Feindt, P., I. Schreiner et J. Bushery (1997). Reinterview: a tool for survey quality management. Dans *Proceedings of the Section on Survey Research Methods*. Alexandria, Virginie, American Statistical Association, pp. 105-110.
- Forsman, G., et I. Schreiner (1991). The design and analysis of reinterview: an overview. Dans *Measurement Errors in Surveys*, P. Biemer et al., eds. New York, John Wiley and Sons, pp. 279-302.
- Fowler, F. J. (1991). Reducing interviewer-related error through interviewer training, supervision and other means. Dans *Measurement Errors in Surveys*, P. Biemer et al., eds. New York, John Wiley and Sons, pp. 259-275.
- Groves, R. M. (1989). *Survey Errors and Survey Costs*. New York, John Wiley and Sons.
- _____, et L. J. Magilavy (1986). Measuring and explaining interviewer effects. *Public Opinion Quarterly*, vol. 50, pp. 251-256.
- Hastie, R., et D. Carlston (1980). Theoretical issues in person memory. In *Person Memory: The Cognitive Basis of Social Perception*, R. Hastie et al., eds. Hillsdale, New Jersey, Lawrence Erlbaum, pp. 1-53.

- Hill, D. H. (1994). The relative empirical validity of dependent and independent data collection in a panel survey. *Journal of Official Statistics*, vol. 10, n° 4, pp. 359-380.
- Huang, H. (1993). *Report on SIPP Recall Length Study*. Internal United States Bureau of the Census, Washington, D.C.
- Institute for Resource Development/Macro Systems, Inc. (1990). *An Assessment of DHS 1 Data Quality*. Demographic and Health Surveys Methodological Reports, n° 1. Columbia, Maryland, Institute for Resource Development/Macro Systems, Inc.
- Jenkins, C., et D. Dillman (1997). Towards a theory of self-administered questionnaire design. Dans *Survey Measurement and Process Quality*, L. Lyberg *et al.*, eds. New York, John Wiley and Sons, pp. 165-196.
- Kahn, R. L., et C. F. Cannell (1957). *The Dynamics of Interviewing*. New York, John Wiley and Sons.
- Kalton, G., D. Kasprzyk et D. B. McMillen (1989). Non-sampling errors in panel surveys. Dans *Panel Surveys*, D. Kasprzyk *et al.*, eds. New York, John Wiley and Sons, pp. 249-270.
- Kalton, G., D. McMillen et D. Kasprzyk (1986). Non-sampling error issues in SIPP. Dans *Proceedings of the Bureau of the Census Second Annual Research Conference*. Washington, D.C., pp. 147-164.
- Kantorowitz, M. (1992). *Methodological Issues in Family Expenditure Surveys*, Vitoria-Gastiers, autonomous community of Euskadi, Euskal Estatistika-Erakundea, Instituto Vasco de Estadística.
- Kish, L. (1962). Studies of interviewer variance for attitudinal variables. *Journal of the American Statistical Association*, vol. 57, pp. 92-115.
- Lyberg, L., et D. Kasprzyk (1991). Data Collection Methods and Measurement Errors: An Overview. dans *Measurement Errors in Surveys*, P. Biemer *et al.*, eds. New York, John Wiley and Sons, pp. 237-258.
- _____, P. Biemer, M. Collins, E. D. DeLeeuw, C. Dippo, N. Schwartz et D. Trewin (1997). Dans *Survey Measurement and Process Quality*. New York, John Wiley and Sons.
- Marquis, K. H., et C. F. Cannell (1971). Effects of some experimental techniques on reporting in the health interview. Dans *Vital and Health Statistics*, Washington, D.C., Public Health Service, Séries 2 (Data Evaluation and Methods Research), n° 41.
- _____, et J. C. Moore (1990). Measurement errors in SIPP program reports. In *Proceedings of the Bureau of the Census 1990 Annual Research Conference*. Washington, D.C., pp. 721-745.
- Mathiowetz, N. (2000). The effect of length of recall on the quality of survey data. Dans *Proceedings of the 4th International Conference on Methodological Issues in Official Statistics*. Stockholm, Statistics Sweden. Peut être consulté sur : http://www.scb.se/Grupp/Omscb/_Dokument/Mathiowetz.pdf (accès le 3 juin 2004).
- _____, et K. McGonagle (2000). An assessment of the current state of dependent interviewing in household surveys. *Journal of Official Statistics*, vol. 16, pp. 401-418.
- Nations Unies (1982). *National Household Survey Capability Programme: Non-sampling Errors in Household Surveys: Sources, Assessment, and Control: Preliminary Version*. DP/UN/INT-81-041/2. New York, Nations Unies, Département de la coopération technique pour le développement et Bureau de statistique.
- Neter, J. (1970). Measurement errors in reports of consumer expenditures. *Journal of Marketing Research*, vol. VII, pp. 11-25.
- _____, et J. Waksberg (1964). A study of response errors in expenditure data from household interviews. *Journal of the American Statistical Association*, vol. 59, pp. 8-55.

- Nolin, M. J., et K. Chandler (1996). *Use of Cognitive Laboratories and Recorded Interviews in the National Household Education Survey*. Washington, D.C., United States Department of Education, National Center for Education Statistics. NCES 96-332.
- Oksenberg, L., C. Cannell et S. Blixt (1996). Analysis of interviewer and respondent behavior in the household survey. *National Medical Expenditure Survey Methods*, 7. Rockville, Maryland, Agency for Health Care and Policy Research, Public Health Service.
- O'Muircheartaigh, C. (1982). *Methodology of the Response Errors Project*. WFS Scientific Reports, n° 28. Voorburg, Pays-Bas, International Statistical Institute.
- _____ (1984a). *The Magnitude and Pattern of Response Variance in the Lesotho Fertility Survey*. WFS Scientific Reports, n° 70. Voorburg, Pays-Bas, International Statistical Institute.
- _____ (1984b). *The Magnitude and Pattern of Response Variance in the Peru Fertility Survey*. WFS Scientific Reports, n° 45. Voorburg, Pays-Bas, International Statistical Institute.
- Schreiner, I., K. Pennie et J. Newbrough (1988). Interviewer falsification in Census Bureau Surveys. Dans *Proceedings of the Section on Survey Research Methods*. Alexandria, Virginie, American Statistical Association, pp. 491-496.
- Schuman, H. et S. Presser (1981). *Questions and Answers in Attitude Surveys*. New York, Academic Press.
- Schwarz, N. (1997). Questionnaire design: the rocky road from concepts to answers. Dans *Survey Measurement and Process Quality*, L. Lyberg *et al.*, eds. New York, John Wiley and Sons, pp. 29-46.
- _____, R. M. Groves et H. Schuman (1995). *Survey Methods*. Survey Methodology Program Working Paper Series. Ann Arbor, Michigan, Institute for Survey Research, University of Michigan.
- _____, et H. Hippler (1991). Response alternatives: the impact of their choice and presentation order. In *Measurement Errors in Surveys*, P. Biemer *et al.*, eds. New York, John Wiley and Sons, pp. 41-56.
- _____, et S. Sudman (1996). *Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research*. San Francisco, Californie, Jossey-Bass.
- Silberstein, A., et S. Scott (1991). Expenditure diary surveys and their associated errors. Dans *Measurement Errors in Surveys*, P. Biemer *et al.*, eds. New York, John Wiley and Sons, pp. 303-326.
- Singh, S. (1987). Evaluation of data quality. Dans *The World Fertility Survey: An Assessment*, J. Cleland et C. Scott, eds. New York, Oxford University Press, pp. 618-643.
- Sirken, M. *et al.* (1999). *Cognition and Survey Research*. New York, John Wiley and Sons.
- Stanton, C., N. Abderrahim et K. Hill (1997). *DHS Maternal Mortality Indicators: An Assessment of Data Quality and Implications for Data Use*. Demographic and Health Surveys Analytical Report, n° 4. Calverton, Maryland, Macro International, Inc.
- Sudman, S., N. Bradburn et N. Schwarz (1996). *Thinking about Answers: The Application of Cognitive Processes to Survey Methodology*. San Francisco, Californie, Jossey-Bass.
- _____, *et al.* (1977). Modest expectations: the effect of interviewers' prior expectations on response. *Sociological Methods and Research*, vol. 6, n° 2, pp. 171-182.
- Tucker, C. (1997). Methodological issues surrounding the application of cognitive psychology in survey research. *Bulletin of Sociological Methodology*, vol. 55, pp. 67-92.
- United States Bureau of the Census (1985). *Evaluating Censuses of Population and Housing*. Statistical Training Document. Washington, D.C. ISP-TR-5.

- _____ (1998). *Survey of Income and Program Participation (SIPP) Quality Profile*, 3^e éd. Washington, D.C., United States Department of Commerce.
- United States Federal Committee on Statistical Methodology (2001). *Measuring and Reporting Sources of Error in Surveys*, Statistical Policy working Paper, No. 31. Washington, D.C., United States Office of Management and Budget. Peut être consulté sur : <http://www.fcsm.gov> (accès le 14 mai 2004).
- United States Law Enforcement Assistance Administration (1972). *San Jose Methods Test of Known Crime Victims*. Statistics Technical Report No.1. Washington, D.C.
- Vaessen, M., et al. (1987). Translation of questionnaires into local languages. Dans *The World Fertility Survey: An Assessment*, J. Cleland et C. Scott, eds. New York, Oxford University Press, pp. 173-191.
- Weiss, C. (1968), Validity of welfare mothers' interview response. *Public Opinion Quarterly*, vol. 32, pp. 622-633.
- Westoff, C., N. Goldman et L. Moreno (1990). *Dominican Republic Experimental Study: An Evaluation of Fertility and Child Health Information*. Princeton, New Jersey, Office of Population Research; et Columbia, Maryland, Institute for Resource Development/Macro Systems, Inc.
- Willis, G. B. (1994). *Cognitive Interviewing and Questionnaire Design; A Training Manual*. Cognitive Methods Staff Working Paper, No. 7. Hyattsville, Maryland, National Center for Health Statistics.
- _____, P. Royston et D. Bercini (1991). The use of verbal report methods in the development and testing of survey questions. Dans *Applied Cognitive Psychology*, vol. 5, pp. 251-267.
- Woltman, H. F., et J. B. Bushery (1977). *Update of the National Crime Survey Panel Bias Study*. Internal United States Bureau of the Census report, Washington, D.C.

Chapitre X

L'assurance de qualité dans les enquêtes : normes, directives et procédures

T. BEDIRHAN ÜSTUN, SOMNATH CHATTERJI,
ABDELHAY MECHBAL ET CHRISTOPHER J. L. MURRAY

Au nom des collaborateurs à l'Enquête sur la santé dans le monde (ESM)*
Organisation mondiale de la santé
Bases factuelles et Information à l'appui des politiques
Genève, Suisse

* La liste des collaborateurs à l'ESM figure sur le site Internet : (<http://www.who.int/whs/>).

RÉSUMÉ

La qualité d'une enquête est de la plus haute importance pour l'exactitude, la fiabilité et la validité de ses résultats. Les équipes chargées d'enquêtes doivent systématiquement mettre en œuvre des procédures d'assurance de qualité afin d'éviter le recours à des pratiques inacceptables et de limiter le plus possible des erreurs lors de la collecte de données. La mise en place de stratégies efficaces et efficientes en vue de l'amélioration de la qualité d'une enquête permet d'assurer la collecte en temps voulu de données de qualité et la validité des résultats. « L'assurance de qualité » peut être également perçue comme un outil d'organisation pour l'application de normes opérationnelles préétablies concernant la structure, le processus et les résultats de l'enquête. Les équipes chargées de l'enquête doivent respecter des normes précises de qualité et suivre les procédures prescrites pour atteindre le niveau de qualité requis. Ces procédures doivent être transparentes, et leur application doit être suivie de façon systématique et confirmée avec soin dans le rapport général sur l'exécution de l'enquête et ses résultats. Il importe également que ces résultats soient mesurés et récapitulés dans toute la mesure du possible à l'aide d'indicateurs chiffrables.

Le présent chapitre décrit une approche systématique employée pour effectuer des mesures d'assurance de qualité allant au-delà des simples mécanismes de contrôle. Une grande enquête internationale – l'Enquête sur la santé dans le monde (ESM) exécutée par de multiples organismes d'enquête dans 71 pays – est utilisée pour illustrer l'élaboration de l'application d'un programme complet d'assurance de qualité. Cette enquête avait été conçue pour recueillir des données comparables afin d'évaluer les différentes dimensions des systèmes de santé dans les pays participants en utilisant des échantillons représentatifs de ces pays. Compte tenu de l'importance des résultats de l'ESM, des procédures rigoureuses d'assurance de qualité ont été mises en place et des experts internationaux ont été utilisés pour faire fonction de groupe d'évaluation externe par des pairs et pour aider les pays à appliquer les normes de qualité convenues d'un commun accord et applicables pour des questions telles que : la méthodologie de sélection des échantillons, l'obtention de taux de réponse acceptables, le traitement des données manquantes, le calcul des mesures de

fiabilité et les vérifications de la comparabilité des données entre les sous-groupes de population et entre les pays.

Termes clés : assurance de qualité, indicateurs de qualité, Enquête sur la santé dans le monde, données manquantes, taux de réponse, échantillonnage, fiabilité, comparabilité au sein d'une population, comparaisons internationales.

A. Introduction

1. L'un des traits essentiels de la conception et de l'exécution d'une enquête est sa « qualité » (Lybert *et al.*, 1997). Dans toute collecte de données, les résultats dépendent des moyens mis en œuvre. Outre la qualité de ses instruments et de ses techniques analytiques, celle de ses résultats dépend principalement de sa mise en œuvre, c'est-à-dire de l'application de saines méthodes d'échantillonnage et de la bonne utilisation du questionnaire.

2. Pour obtenir la meilleure qualité, chaque équipe de l'enquête doit se conformer à un ensemble type de directives d'exécution. Ces directives portent sur les éléments suivants :

- a) Des normes de qualité qu'il convient de respecter à chaque étape de l'enquête;
- b) De procédures d'assurance de qualité [*Quality assurance (QA)*] qui définissent les mesures à prendre pour suivre le déroulement de l'enquête sur le terrain;
- c) L'évaluation d'un processus d'assurance de qualité qui mesure l'impact des normes d'assurance de qualité sur les résultats de l'enquête et de procédures tendant à améliorer la pertinence et l'efficacité de l'ensemble du processus d'assurance de qualité (Biemer *et al.*, 1991).

3. Les directives ont pour objectif général d'aider à améliorer la qualité plutôt que d'assurer l'audit de l'exécution de l'enquête. Comme toute enquête représente un gros investissement qui fait intervenir différentes parties et dont les résultats influent sur les politiques du pays, il est essentiel de traiter la qualité avec beaucoup de sérieux. L'assurance de qualité est perçue comme un processus continu tout au long de l'enquête, depuis la préparation et l'échantillonnage jusqu'à la collecte et l'analyse des données et à la rédaction du rapport. Les directives visent aussi à assurer une meilleure compréhension de la conception de l'enquête par les utilisateurs. L'établissement de procédures types a pour but d'assurer que :

- La collecte de données est pertinente et correspond aux besoins du pays;
- Les données peuvent être comparées à l'intérieur d'un même pays et d'un pays à un autre afin de permettre d'identifier les similarités et les différences entre populations;
- L'exécution pratique de l'enquête est conforme à des protocoles acceptés;
- Les erreurs dans la collecte des données sont réduites au minimum;
- La capacité de collecte de données s'améliore peu à peu.

B. Normes et procédures d'assurance de qualité

4. L'assurance de qualité (Statistique Canada, 1998) est définie comme toute méthode ou procédure de collecte, traitement ou analyse des données d'une enquête qui vise à maintenir ou à rehausser leur fiabilité ou leur validité. L'assurance de qualité peut s'interpréter comme ayant des significations similaires et pourtant différentes. Dans le présent chapitre, nous utilisons le paradigme de gestion d'une qualité totale,

qui examine le processus d'enquête à chacune de ses phases et s'efforce d'ébaucher une approche non seulement pour réduire les erreurs d'échantillonnage et autres mais aussi pour améliorer la pertinence et la faisabilité des enquêtes ainsi que la capacité du pays à les effectuer. Pour atteindre ce but tout en restant pratique, ce chapitre utilise les normes et procédures d'assurance de qualité de l'Enquête sur la santé dans le monde (ESM) [Organisation mondiale de la santé, 2002] en se référant à toutes les opérations, y compris :

- La sélection des organismes qui seront chargés de l'enquête;
- L'échantillonnage;
- La traduction;
- La formation;
- L'exécution de l'enquête;
- La saisie et l'enregistrement des données;
- L'analyse des données;
- Les indicateurs de qualité;
- Les rapports par pays;
- Les visites sur place.

5. La figure X.1 décrit l'ensemble du cycle de l'ESM en indiquant les opérations susmentionnées à chaque phase de l'exécution de l'enquête. Les directives d'assurance de qualité qui ont été rédigées par un grand nombre de participants à l'ESM ainsi que par des experts internationaux visent à identifier les pratiques optimales dont l'application, pour l'exécution et le suivi d'une enquête de qualité, est possible. Chaque opération entrant dans l'exécution d'une enquête implique un certain examen de la qualité. Par exemple, il importe que les instruments de l'enquête aient de bonnes propriétés, que l'échantillonnage soit représentatif de la population cible et que les données soient claires et complètes.

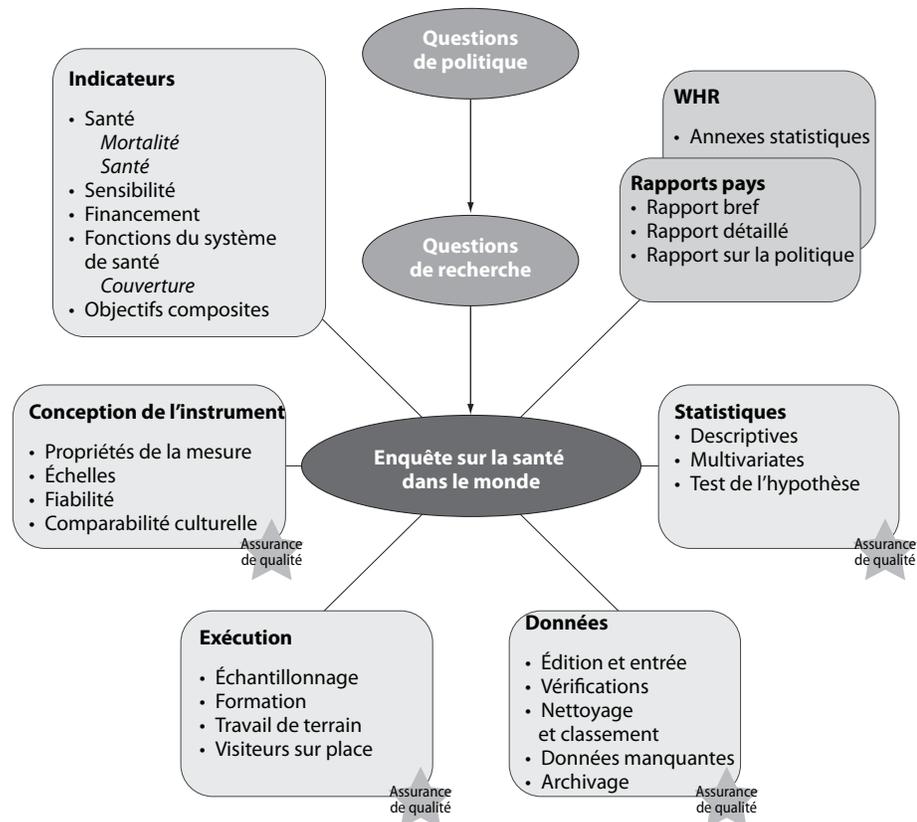
6. Cet ensemble de procédures ne constitue qu'un exemple destiné à présenter « l'approche de l'assurance de qualité » pour la conception et la mise en œuvre d'une enquête comme un processus et pour améliorer les résultats d'une enquête du point de vue de sa pertinence, de son exactitude, de sa cohérence et de sa comparabilité. Toute équipe chargée de concevoir et de mettre en œuvre une enquête peut utiliser une approche similaire en ne perdant pas de vue les buts spécifiques de son enquête et l'applicabilité des normes d'assurance de qualité proposées dans le présent chapitre. Surtout, la qualité doit retenir particulièrement l'attention et doit être suivie dans son contexte opérationnel. Les résultats du processus d'assurance de qualité doivent être indiqués en termes quantitatifs au moyen d'indicateurs appropriés lorsqu'une mesure est possible (par exemple, les taux d'échantillonnage, les taux de réponse, les données manquantes, la fiabilité de tests répétés) et en termes qualitatifs récapitulant la structure, le processus et les résultats de l'enquête.

C. Application pratique des directives d'assurance de qualité : l'exemple des Enquêtes sur la santé dans le monde

7. La stratégie générale d'assurance de qualité décrite ci-dessus a été appliquée lors de l'ESM pour améliorer la qualité des enquêtes, notamment dans plusieurs pays en développement d'Asie et d'Afrique subsaharienne. La présente section vise à utiliser les normes, procédures et rapports d'assurance de qualité comme guide pratique. D'autres équipes peuvent utiliser cet exemple selon leurs besoins. À notre connais-

sance, cet exemple est le premier cas d'application systématique de procédures d'assurance de qualité à des enquêtes internationales, et les organismes d'exécution et leurs collaborateurs les ont trouvées très utiles pour l'organisation de leur travail et l'établissement de leurs rapports. Il ressort des données initiales que l'on a pu détecter très tôt des erreurs et les éliminer, et accroître ainsi l'exactitude et l'efficacité des résultats.

Figure X.1
Procédures d'assurance de qualité de l'ESM



8. L'Organisation mondiale de la santé (OMS) a été l'initiatrice de l'Enquête sur la santé dans le monde, véritable programme de collecte de données réelles permettant de recueillir des informations sur la santé des populations et les systèmes de santé de façon continue (Üstün *et al.*, 2003a, 2003b; Valentine, de Silva et Murray, 2000; Organisation mondiale de la santé, 2000). L'ESM répond au besoin des pays de disposer d'un système détaillé et durable d'informations sur la santé par le biais d'enquêtes permettant de mesurer les paramètres de santé essentiels de la population; il permet également de réunir des procédures et instruments normalisés pour les enquêtes générales sur la population, de manière à présenter des données comparables pour tous les États membres de l'OMS. Ces méthodes et instruments sont de structure modulaire et ont été améliorés par une étude scientifique des écrits, de longues consultations avec des experts internationaux et des essais pilotes à grande échelle effectués dans plus de 63 pays et 40 langues (Üstün *et al.*, 2003a, 2003c; 2001). L'ESM est conçue pour évoluer au fur et à mesure de son application sous l'effet des apports continus de collaborateurs, y compris de décideurs, d'organismes d'enquête, de chercheurs et d'autres parties intéressées. Les pays et l'OMS sont conjointement détenteurs des données, et il y a une volonté de part et d'autre de collecte durable de données, de renforcement des capacités locales et d'utilisation des résultats de l'enquête pour guider le développement et la mise en œuvre de la politique de santé.

9. Ce chapitre examine systématiquement chaque étape du processus d'enquête, sauf la conception du questionnaire et la façon dont il est testé, qui sont examinés ailleurs (voir Üstün *et al.*, 2003b), et présente les normes d'assurance de qualité de l'ESM pour chacune d'entre elles. Ces normes doivent permettre d'améliorer l'efficacité et d'empêcher le recours à des pratiques inacceptables. La qualité mérite une attention d'autant plus grande aujourd'hui que les données de l'ESM revêtent une importance croissante pour les États membres de l'OMS et influent de plus en plus sur les politiques de santé. C'est pourquoi des directives générales ont été formulées pour donner plus de fiabilité et de validité aux enquêtes de l'ESM en réduisant la fréquence des erreurs qui peuvent être évitées. Les directives d'assurance de qualité ainsi adoptées deviendront des outils de première importance pour l'organisation de l'ESM ainsi que pour la préparation et la conduite des enquêtes. Ce chapitre offre donc un guide général pour les aspects critiques qui réclament une attention particulière, notamment pour la collecte de données de qualité.

10. Ces directives serviront également de canevas d'évaluation aux responsables des enquêtes et aux spécialistes de l'assurance de qualité (réseau d'experts internationaux ayant une vaste expérience des enquêtes chargés d'évaluer l'ensemble du processus). Ces personnes effectueront des visites sur le terrain pour aider à la mise en œuvre de l'ESM et effectuer une évaluation structurée et approfondie du processus, aidant ainsi les pays à procéder à une évaluation systématique de la qualité de leurs enquêtes et à identifier les domaines d'activité qui pourraient être améliorés.

1. Sélection des organismes d'enquête

11. L'exécution d'une enquête nationale nécessite beaucoup de connaissances, d'aptitudes, de ressources et de compétences. C'est ce qui explique que les enquêtes sont organisées selon des styles et des traditions qui diffèrent selon les pays et selon les secteurs. Pour faire en sorte que l'ESM soit confiée à une équipe compétente dans chaque pays, il importe d'identifier les organismes d'enquête de qualité et de leur imposer des normes comme conditions contractuelles. La pratique adoptée pour l'ESM consiste à consulter les ministères de la santé, les bureaux régionaux et les représentants résidents de l'OMS ou les chargés de liaison afin d'identifier ces organismes. Étant donné la taille et la complexité de ces enquêtes, il convient de démontrer leur faisabilité à travers un processus d'appel d'offres, conformément aux règlements de l'OMS. Ce processus commence par une invitation lancée aux organismes d'enquête compétents afin qu'ils soumettent une offre pour l'ESM, conformément aux spécifications techniques de l'échantillonnage, de la conduite d'entrevues et de la collecte de données. [Les spécifications techniques concernant l'ESM peuvent être consultées sur le site Internet de l'ESM (www.who.int/whs)]. Ces soumissions sont comparées en regard d'un certain nombre de critères avant qu'un organisme soit sélectionné.

12. Parmi les critères d'évaluation des qualifications des organismes candidats figurent les suivants :

- Leurs antécédents (c'est-à-dire leur expérience récente d'au moins cinq grandes enquêtes nationales sur des échantillons d'au moins 3 000 unités);
- Leur aptitude à assurer l'ensemble du processus (à savoir échantillonnage, formation, collecte et analyse des données);
- Leur expérience de divers modes de collecte de données, y compris les entrevues face à face (et d'autres modes possibles, tels que le téléphone, le courrier, l'ordinateur, etc.);

- L'existence de documents sur des enquêtes précédentes (portant notamment sur une évaluation chiffrée de la représentation de l'échantillon, la couverture de la population du pays, la qualité des entrevues, le type de formation et son coût, l'assurance de qualité et d'autres procédures);
- Une documentation sur les échéanciers habituels des enquêtes et l'aptitude de l'organisme à exécuter celles-ci dans les délais prévus;
- L'aptitude de l'organisme à mettre en place et à utiliser une bonne infrastructure concernant les systèmes d'information sur la santé, en travaillant étroitement avec le ministère de la santé, les organismes nationaux de statistiques et d'autres institutions.

13. L'appel d'offres est utile en ce sens qu'il permet d'identifier la meilleure offre possible des points de vue qualité et prix et de faire une comparaison entre toutes les prestataires possibles du pays. L'OMS et le ministère de la santé peuvent ainsi identifier le meilleur organisme d'enquête possible, renforcer sa capacité d'exécution de nouvelles enquêtes et incorporer les données de l'ESM au système d'information sur la santé. Le processus de passation de marché prévoit également des sanctions si les résultats ou leur qualité ne sont pas à la hauteur de ce qui avait été convenu. Les marchés passés avec des consortiums sont à encourager comme moyen d'assurer que des partenaires à la hauteur de la tâche (par exemple, le ministère de la santé en association avec l'office national de statistique) collaborent en vue d'assurer un accès à un bon cadre d'échantillonnage.

14. Les différentes propositions reçues doivent être examinées avec soin sur la base des critères énumérés ci-dessus, et leur analyse comparative doit être documentée.

15. En résumé, il importe non seulement d'identifier un organisme compétent, susceptible d'effectuer l'enquête conformément à ses spécifications techniques, mais aussi d'apporter à celui-ci l'appui technique nécessaire à l'obtention du résultat souhaité. Pour les grandes enquêtes nationales, il est souvent nécessaire, à l'intérieur du pays, de créer un partenariat de groupes, d'institutions et de personnes possédant les compétences nécessaires dans les domaines de la conception, de la formation, de la mise en œuvre, du traitement et de l'analyse des données et de la rédaction de rapports.

2. Échantillonnage

16. L'enquête ne vaut que ce que vaut l'échantillon. Si la conception et/ou l'utilisation de l'échantillon est défectueuse, on ne peut faire grand-chose pour pallier la représentativité limitée de l'échantillon ou pour combler les lacunes laissées par les informations manquantes. Les résultats de l'enquête sont alors faussés dans un sens et dans des proportions que l'on ignore.

17. En raison du large éventail d'applications possibles sur le terrain, l'OMS et un groupe d'experts techniques internationaux ont défini un ensemble de directives afin d'établir un bon échantillon pour l'ESM [les directives d'échantillonnage à l'intention des pays participants peuvent être consultées sur le site Internet de l'OMS (www.who.int/whs)]. Les normes d'échantillonnage scientifique sont fondées sur les méthodes de sélection probabiliste et sont amplement connues et acceptées (Üstün *et al.*, 2001; Kish, 1995a). Toutefois, ces méthodes ne sont généralement pas retenues en raison de leur faible efficacité opérationnelle, du manque de supervision de leur application des procédures d'échantillonnage sur le terrain et/ou de leurs coûts élevés d'application sous certaines conditions et dans certains contextes particuliers.

18. Les directives de l'OMS insistent sur les principes scientifiques de l'échantillonnage en tant que normes précises de qualité, donnent des exemples de bons plans d'échantillonnage et identifient des normes d'assurance de qualité à respecter par les pays. Le cas échéant, l'OMS et les conseillers techniques apportent un appui technique aux pays. Certains aspects importants de l'échantillonnage effectué dans le cadre de l'ESM sont décrits ci-après :

- a) L'échantillon de l'ESM doit cibler la population de fait (c'est-à-dire toutes les personnes vivant dans le pays, y compris les travailleurs immigrés temporaires, les immigrants et les réfugiés) et non pas la population de droit (les seuls ressortissants du pays). Il importe de créer une bonne représentation « en miniature » de la population d'ensemble du pays. Pour ce faire, il est essentiel de représenter toutes les personnes vivant dans le pays et d'assurer la couverture géographique intégrale de celui-ci;
- b) La taille de l'échantillon doit être suffisante pour permettre de bonnes estimations (solides) des grandeurs étudiées aux niveaux national et sous-national selon les objectifs de l'enquête; en même temps, les responsables de l'enquête doivent trouver le juste équilibre entre la nécessité de disposer de plus grands échantillons pour permettre de meilleures estimations et la nécessité de limiter l'accroissement correspondant des coûts de l'enquête. De grands échantillons ne sauraient compenser la mauvaise qualité. Pour diverses raisons, il peut être nécessaire de disposer d'une représentation adéquate des minorités (par exemple, de certains groupes ethniques ou autres), qui peuvent nécessiter un suréchantillonnage (c'est-à-dire une plus forte probabilité de sélection). Si un sous-groupe de population a besoin d'un suréchantillonnage en raison d'une question liée à une étude scientifique, cela doit être précisé de façon détaillée dans les spécifications. En pareil cas, il convient d'appliquer des pondérations différentielles lors de l'analyse des données afin de corriger la distorsion causée par le suréchantillonnage;
- c) Dans l'ESM, un cadre d'échantillonnage (c'est-à-dire une liste des zones géographiques, des ménages ou des individus parmi lesquels l'échantillon est sélectionné, qui pourrait être tirée d'une liste informatisée de la population, d'un recensement récent, des listes électorales, etc.) couvrant 90 % des tous les principaux sous-groupes à étudier est jugé acceptable. Les pays doivent utiliser le cadre d'échantillonnage le plus récent dont ils disposent. S'il est vieux de deux ans ou plus, une énumération ou une liste des ménages établie pour actualiser le cadre à l'avant-dernier stade de la sélection est souvent nécessaire. Des méthodes de comptage rapide peuvent être utilisées pour actualiser les mesures de la taille des unités primaires d'échantillonnage avant la sélection; ces méthodes comprennent le comptage dans certains endroits pour lesquels on ne dispose pas d'un cadre à jour en raison d'une cartographie caduque ou pour d'autres raisons. En dehors des méthodes de comptage rapide dans les zones d'échantillonnage choisies, on peut actualiser le cadre en utilisant d'autres sources telles que les adresses postales des bureaux de poste locaux ou les listes de facturation des abonnés aux compagnies d'alimentation en électricité ou en eau. Il est essentiel que la population soit scientifiquement pondérée par rapport au dernier recensement;
- d) L'échantillon de l'ESM cible tous les membres adultes de la population âgés de 18 ans ou plus¹. Dans la plupart des cas, il utilise le dernier recensement comme cadre d'échantillonnage. Les ménages sont choisis selon une procé-

¹ Actuellement, l'ESM n'inclut que les adultes. Pour l'avenir, il est également prévu d'effectuer une enquête qui comprendra également les enfants.

de d'échantillonnage stratifié en grappes à plusieurs degrés. Un individu par ménage est ensuite sélectionné selon une procédure de sélection au hasard [par exemple, la méthode de la table de Kish (Kish, 1995a), ou d'autres méthodes telles que celle du dernier anniversaire, et la méthode Trohdahl/Carter/Bryant (Bryant, 1975)]. On peut aussi, à ce stade, utiliser des tables de nombres aléatoires, à condition que les nombres choisis soient soigneusement documentés. Quelle que soit la technique de sélection utilisée, il faut s'efforcer le plus possible de réduire la distorsion liée à la sélection lors de l'exécution sur le terrain. Les pays peuvent chercher à concevoir le plan d'échantillonnage le plus simple possible qui réponde aux objectifs de mesure de l'enquête. Dans le cas d'une conception d'une complexité excessive, l'exécution peut être difficile et engendrer un nombre d'erreurs qui rendent l'échantillon inutilisable. La faisabilité et la nécessité de pouvoir suivre la conception de l'échantillon sont deux éléments clés pour la qualité;

² Les Nations Unies définissent le ménage comme un groupe de personnes qui vivent sous le même toit et partagent la même cuisine et la même table (autrement dit, dont la nourriture provient de la même source). Pour l'ESM, une personne est généralement considérée comme faisant partie du ménage si elle se trouve actuellement dans une institution en raison de son état de santé, et doit donc être incluse dans la liste des membres du ménage.

- e) L'ESM utilise la définition du ménage donnée par les Nations Unies²; toutefois, il peut se produire certaines variations dans cette définition selon les circonstances locales. L'incidence possible de ces variations sur l'échantillonnage devrait être indiquée dans les rapports des pays. Si les pays utilisent un cadre d'échantillonnage des ménages, il est donc suggéré qu'ils utilisent la même définition du ménage pour l'enquête que celle qu'ils ont utilisée pour le cadre initial;
- f) L'ESM utilise une stratégie d'échantillonnage scientifique, qui comporte une probabilité de sélection non nulle pour tout individu inclus dans l'enquête. L'utilisation de strictes méthodes probabilistes à chaque stade de l'échantillonnage est cruciale et permet d'extrapoler les données de l'échantillon à l'ensemble de la population. Autrement, les résultats de l'enquête ne seraient ni représentatifs ni valides;
- g) Il est difficile d'inclure les personnes se trouvant dans des institutions dans une enquête générale sur la population car il faut établir pour elles des cadres distincts. Le fait d'interroger les personnes qui se trouvent dans des hôpitaux, des maisons de retraite, des casernes ou des prisons comporte également un certain nombre de problèmes moraux. Étant donné les différences énormes qui existent entre ces institutions d'un pays à un autre, il n'y a pas de solution universelle. L'ESM inclut parfois les personnes hospitalisées s'il est possible de les interroger durant l'enquête. On utilise ensuite le pourcentage de la population se trouvant en institution tel qu'il ressort du recensement pour vérifier la concordance des taux qui se dégagent de l'enquête. Cela est particulièrement important pour l'ESM car les personnes qui se trouvent dans des institutions telles que maisons de retraite ou qui font de longs séjours à l'hôpital risquent fort d'être en moins bonne santé que le reste de la population et doivent donc être incluses dans l'échantillon pour réduire le risque de sous-estimation de l'état de santé;
- h) Les directives d'échantillonnage de l'ESM expliquent clairement ce que l'on entend par non-réponse unitaire et calcul des taux de non-réponse vis-à-vis des échantillons. La stratégie d'échantillonnage de l'ESM ne permet pas de substituer un autre ménage ou un autre individu à un cas de non-réponse;
- i) Les résultats de l'enquête en ce qui concerne l'échantillonnage doivent indiquer les erreurs types pour les variables importantes de l'enquête, afin que les utilisateurs puissent voir l'erreur de mesure en termes statistiques;

- j) L'utilisation des Systèmes d'information géographique [*Geographical Information Systems* (GIS)] peut aider à améliorer la qualité des résultats en permettant de vérifier l'application sur le terrain du plan d'échantillonnage; autrement dit, de vérifier que les entrevues ont bien eu lieu à un certain endroit et qu'elles n'ont pas été inventées (De Lepper, Scholten and Stern, 1995). Le GIS peut aussi donner une valeur supplémentaire aux données en liant les informations telles que la distance qui sépare le ménage des établissements de soins, des ressources en eau et d'autres ressources environnementales aux paramètres de santé mesurés (état de santé, maladies, facteurs de risque) dans l'enquête. Il peut aussi montrer sur une carte les caractéristiques de dispersion d'un paramètre donné, et indiquer ainsi les inégalités en matière de santé. Pour ce faire, l'ESM utilise le Système mondial de positionnement (GPS) et des cartes numérisées pour géocoder les données selon certaines directives (voir <http://www3.who.int/whosis/gis>). Certaines dispositions légales ont été prises pour maintenir la confidentialité des informations personnelles car l'information géocodée risque de violer les normes de protection.

Évaluation de l'échantillonnage

19. Il convient d'évaluer la stratégie d'échantillonnage avant le démarrage de l'enquête afin de déterminer le bien-fondé de la stratification, l'adéquation de la représentation de la population et la taille et la réparation des grappes sélectionnées. Le rapport doit documenter soigneusement les procédures utilisées sur le terrain et noter tout écart par rapport à la conception afin que les utilisateurs puissent mieux se faire une idée de la qualité des résultats de l'enquête.

20. Lors de la collecte de données, il faut suivre de près la sélection des ménages et des individus par le personnel de terrain et/ou les superviseurs travaillant dans les bureaux afin de s'assurer de l'exactitude, par exemple, dans l'utilisation des tables de Kish et l'établissement des listes de ménages.

21. Après la collecte des données, l'analyse chiffrée des données (discutée ci-après) est utilisée pour évaluer la qualité des données au moyen :

- D'une statistique sommaire, que nous appelons « l'indice de déviation de l'échantillon » [*sampling deviation index* (SDI)];
- De tests répétés de la fiabilité, pour indiquer la « stabilité » de l'instrument dans son utilisation par divers enquêteurs;
- D'informations sur la quantité de cas de non-réponse et de données manquantes.

22. Ces procédures sont décrites de façon plus détaillée dans la section sur l'analyse des données. Le tableau X.1 présente une liste sommaire des déterminants de la qualité de l'échantillonnage.

Tableau X.1

Liste sommaire des déterminants de la qualité de l'échantillonnage

- Aperçu général de la composition de la population (urbaine/rurale, minorités, langues, groupes suréchantillonnés)
- Cadre d'échantillonnage et nombre des phases de l'échantillonnage :
- Le(s) cadre(s) d'échantillonnage couvre-t-il (couvrent-ils) l'ensemble de la population cible ?
- À quand remonte le cadre d'échantillonnage ?
- Stratification au sein du cadre d'échantillonnage
- Unités d'échantillonnage à chaque stade : probabilité de sélection connue
- Taille des unités d'échantillonnage à chaque stade : veiller à ce que toutes les unités d'échantillonnage soit d'une taille qui dépasse un minimum prédéterminé
- Vérifier « sur le terrain » la taille des unités et certaines questions telles que celle de savoir s'il y a un ou plusieurs ménages à « l'adresse » sélectionnée, et comment choisir parmi ceux-ci
- Taille de l'échantillon sélectionné
- Pondération de probabilité pour le ménage
- Pondération de probabilité pour le répondant
- Formation à l'utilisation et à la bonne application de la table de Kish (ou d'un autre instrument)
- Vérification de la procédure de sélection du répondant au sein des ménages
- Rapport sommaire sur l'échantillonnage, sa mise en œuvre, les déviations, les pondérations, les erreurs types

3. Traduction

23. Pour pouvoir faire des comparaisons significatives des données entre différentes cultures, il faut disposer d'un instrument pertinent qui mesure les mêmes paramètres dans différents pays. L'instrument de l'ESM a été mis au point au terme d'une étude scientifique des instruments d'enquête existants, de consultations approfondies avec des experts et d'essais systématiques sur le terrain lors d'une enquête effectuée sur plusieurs pays (Üstün *et al.*, 2003a). Nous avons indiqué les caractéristiques des instruments de l'enquête, leur pertinence et leur applicabilité culturelle ailleurs (Üstün *et al.*, 2003b). Pour toute autre enquête, les concepteurs doivent chercher à avoir les meilleurs instruments et mesures et s'assurer que ces instruments conviennent à leurs besoins, sont précis dans leurs mesures et ont passé avec succès des essais pilotes pour tester leur faisabilité et leur stabilité.

24. Une fois que l'on dispose d'un bon instrument, la traduction est l'un des moyens essentiels de garantir que les questions ont la même signification dans les différentes langues. Étant donné les sociétés multiraciales dans lesquelles nous vivons, il est essentiel de disposer de bonnes traductions qui permettent de mesurer les mêmes concepts de l'enquête.

25. Souvent, dans un même pays, l'instrument doit être traduit dans plusieurs langues selon l'ampleur des différents groupes linguistiques du pays. Il est suggéré que chaque groupe linguistique qui représente 5 % de la population soit interrogé dans sa propre langue. Pour les répondants qui sont interrogés dans une langue dans laquelle

il n'existe pas de version du questionnaire, il faut s'assurer de la bonne compréhension des principaux concepts. Les enquêteurs travaillent avec l'une des traductions existant dans le pays pour poser des questions dans la langue dans laquelle il n'y a pas de traduction, utilisant pour cela les directives générales. L'autre difficulté à laquelle on se heurte lors d'une grande enquête effectuée dans plusieurs pays tient au fait que, par exemple, dans beaucoup de pays d'Afrique et d'Asie, les langues ne sont pas écrites et l'on ne peut disposer de questionnaires écrits. En pareils cas, il est recommandé qu'une traduction type soit faite dans une langue connue dans le pays conformément aux directives et que l'on en établisse une version écrite par translittération.

26. Les directives pour la traduction des instruments de l'ESM sont le fruit de la longue expérience que possède l'OMS de la conception et de la mise en œuvre d'études internationales avec de multiples partenaires et experts linguistiques. Les directives de traduction de l'ESM, qui peuvent être consultées sur le site Internet de l'ESM (www.who.int/whs), soulignent combien il importe de maintenir l'équivalence des concepts et de suivre une procédure qui identifie les pièges possibles et évite les contresens. Ces directives précisent que :

- La traduction doit tendre à produire un questionnaire compréhensible à l'échelon local;
- L'intention initiale des questions doit être exprimée de la meilleure façon possible dans la langue locale;
- Les spécifications doivent tendre à exprimer la signification originale des questions et des différentes réponses possibles précodées;
- Le questionnaire devrait être traduit d'abord par des spécialistes de la santé et des enquêtes ayant une bonne compréhension des principaux concepts traités. Un ensemble des termes clés et des termes jugés problématiques lors de la première traduction directe devraient être retraduits par des experts linguistes qui expliqueraient ensuite toutes les interprétations possibles de ces termes et suggéreraient d'autres formulations. Un groupe de rédaction sous la supervision du principal responsable national de l'enquête examinerait ensuite la traduction et la retraduction et rendrait compte à l'OMS de la qualité de la traduction;
- Pour améliorer la qualité, il conviendrait de faire appel à des groupes de réflexion et d'utiliser des méthodes linguistiques qualitatives telles que l'établissement d'un inventaire d'expressions locales et la comparaison d'expressions à celles d'autres langues. L'OMS a déjà entrepris des études systématiques de la traduction et du contrôle des connaissances en certaines langues et tenu compte des résultats de ces études dans le texte actuel du questionnaire de l'ESM. Il est encore recommandé d'effectuer un « contrôle des connaissances » (ou plus précisément la vérification des interprétations données aux questions) à l'aide du questionnaire traduit auprès d'éléments de la population locale. Tous les documents relatifs à l'ESM (à savoir, le questionnaire, les spécifications question par question, le manuel de l'enquête et les manuels de formation) doivent obligatoirement être traduits dans la langue locale. Le programme d'entrée des données peut rester en anglais. Toutefois, si le pays a traduit le questionnaire par des moyens électroniques en suivant les spécifications de l'OMS, ce programme peut être traduit automatiquement dans les autres langues;
- Chaque pays de l'ESM doit soumettre un rapport sur la qualité du travail de traduction à la fin de la phase pilote. Pour les points qui ont été jugés particulièrement difficiles à traduire, il conviendrait de demander des formulaires d'évaluation linguistique décrivant la nature de la difficulté de traduction.

- Les conseillers du pays en assurance de qualité devraient accorder une attention particulière aux opérations de mise en œuvre dans le processus de traduction et devraient vérifier la liste des termes clés avec le principal responsable national de l'enquête;
- Dans les pays où sont parlés plusieurs dialectes et/ou langues qui ne peuvent être transcrits par écrit, des protocoles particuliers de traduction devraient être discutés avec l'OMS.

Évaluation de la traduction

27. La traduction complète du questionnaire doit être soumise à l'OMS avec le démarrage des entrevues pilotes de l'ESM. Cette traduction devrait être vérifiée par des experts compétents dans les langues en question, qui devraient, s'il y a lieu, faire part de leurs commentaires au pays.

28. Il conviendrait d'examiner la liste des termes clés retraduits et le rapport sur le processus de traduction et les questions qu'il suscite. Les fiches d'évaluation linguistiques (Üstün *et al.*, 2001) doivent être examinées systématiquement par le responsable national de l'enquête, puis par l'OMS, pour déceler les points particulièrement problématiques et favoriser une solution commune pour toutes les langues chaque fois que cela est possible.

29. Des discussions devraient avoir lieu avec les enquêteurs concernant la compréhension des procédures employées sur le terrain lorsqu'un terme, une expression ou une question n'est pas compris. Lors de ces discussions, il devrait être question de la mesure dans laquelle les enquêteurs sont censés « expliquer » et « interpréter » les questions pour les répondants.

Tableau X.2

Liste récapitulative des procédures d'examen de la traduction

- Langues parlées dans le pays, couverture des principaux groupes linguistiques
- Qui a participé au processus de traduction ?
- Tous les documents nécessaires ont-ils été traduits ?
 - Questionnaire
 - Appendice
 - Guide pour l'administration (seulement lorsque les enquêteurs ne parlent pas anglais)
 - Manuel de l'enquête (seulement lorsque les enquêteurs ne parlent pas anglais)
 - Codes de résultats
- Quels problèmes se sont-ils posés lors de la traduction ?
- Quel protocole a-t-il été retenu (par exemple, envoi de la traduction intégrale à l'OMS ou seulement d'une liste des éléments clés) ?
- Des formulaires d'évaluation linguistique ont-ils été remplis ?

D. Formation

30. La formation de l'équipe chargée de l'enquête est indispensable pour la qualité de celle-ci. La formation est un processus continu qui doit se dérouler avant

et pendant le processus de collecte de données et être suivie d'une réunion de compte rendu après l'achèvement du travail sur le terrain.

31. La formation doit être dispensée à tous les niveaux de l'équipe chargée de l'enquête, des enquêteurs aux instructeurs et aux superviseurs, ainsi qu'à l'équipe centrale chargée de superviser le processus à l'échelon national, afin que toutes les personnes travaillant à l'enquête aient une idée précise de leur influence sur la qualité des données.

32. La formation a pour objectif général :

- D'assurer une utilisation uniforme des documents de l'enquête;
- D'expliquer le but de l'enquête et de son protocole;
- De motiver les enquêteurs;
- D'offrir des suggestions pratiques;
- D'améliorer la qualité d'ensemble des données.

33. Pour atteindre au moins une partie de cet objectif, l'OMS a organisé des ateliers régionaux de formation pour les principaux responsables de l'ESM de tous les pays participants et produit divers matériels didactiques, y compris une bande vidéo de formation et un disque compact traitant de tous les aspects de la formation.

Sélection des enquêteurs

34. Il importe d'utiliser des enquêteurs expérimentés ainsi que des personnes connaissant bien le sujet de l'enquête.

35. Les enquêteurs devraient avoir au moins terminé leur scolarité dans leur pays et parler couramment sa langue principale. Chaque pays doit décider ensuite du niveau d'instruction complémentaire à exiger et procéder à des évaluations formelles avant la sélection.

36. Le soin de décider si les enquêteurs devraient ou non être des agents de santé est laissé à chaque pays. Les caractéristiques des enquêteurs (âge, sexe, instruction, formation professionnelle, situation d'emploi, expérience d'enquêtes précédentes, etc.) devraient être consignées sur une base de données séparée. Ces informations doivent ensuite pouvoir renvoyer aux numéros d'identification des enquêteurs pour chaque questionnaire rempli, afin de permettre d'évaluer la performance de chaque enquêteur.

Durée et contenu de la formation et méthodes employées

37. La formation doit être suffisamment longue pour permettre aux enquêteurs de se familiariser non seulement avec les bonnes techniques d'enquête mais aussi avec le contenu du questionnaire à utiliser. La durée de la formation doit être fonction de l'expérience des enquêteurs.

38. La durée recommandée de la formation pour l'ESM est de trois à cinq jours, trois jours étant suffisants pour les enquêteurs expérimentés qui n'ont besoin de formation que sur le questionnaire. Pour tous les autres, une formation plus longue est jugée préférable.

39. Toute la formation devrait être dispensée autant que possible par la même équipe, afin que la même formation soit dispensée à tous les enquêteurs, que ce soit au cours d'une même session ou de sessions différentes organisées pour différents groupes à des dates et des lieux différents. Pour réduire les coûts et permettre une formation régionale, la formation peut être décentralisée et dispensée en cascade. Toutefois,

les inconvénients d'une formation diluée ou inégale l'emportent sur les avantages économiques de ces formules.

40. Une session de révision est fortement recommandée, s'il est possible de la prévoir à un certain stade pendant la collecte des données, de préférence vers le milieu de cette période, pour revoir divers aspects de cette collecte, en particulier ceux qui se révèlent complexes et difficiles, et les directives qui ne sont pas suffisamment suivies par les enquêteurs. Cette session devrait également être l'occasion de faire le point de ce qui a déjà été accompli. Elle devrait permettre aux superviseurs et à l'équipe centrale de rendre compte aux enquêteurs des aspects positifs du travail qu'ils ont fait jusqu'ici, et aux enquêteurs d'informer les superviseurs et l'équipe centrale des problèmes rencontrés.

41. Les méthodes de formation devraient comprendre le plus possible de reconstitution des rôles joués dans les entrevues (avec un minimum d'une reconstitution par enquêteur). Cette méthode permet de mieux assimiler les techniques de l'entrevue. Pour qu'elle soit efficace, il faut préparer d'avance plusieurs dialogues de façon à illustrer les différentes articulations de l'entrevue, la nature des explications qui sont permises et les problèmes auxquels on s'attend pendant l'entrevue avec des répondants réticents.

42. Outre cette reconstitution des rôles, il faudrait prévoir au moins une occasion, avant d'entreprendre la collecte de données, de mener une entrevue avec un répondant pris à l'extérieur du groupe confié à l'enquêteur. Ces entrevues préliminaires devraient être enregistrées ou filmées le plus souvent possible pour être examinées et analysées lors des séances de formation. Les pays de l'ESM sont encouragés à enregistrer de leur côté ces entrevues. Chaque rôle joué ou chaque entrevue préliminaire devrait être commenté à l'intention des enquêteurs.

43. Des matériels didactiques devraient être fournis à tous les enquêteurs à titre de références. Tout matériel ainsi fourni devrait faire l'objet d'un examen approfondi pendant la formation et, le cas échéant, être traduit dans les langues du pays.

44. La formation devrait porter sur les éléments suivants :

- Les questions administratives;
- La planification du travail de terrain;
- L'examen de tous les matériels fournis;
- Les méthodes de prise de contact, les formulaires de consentement et la confidentialité.

La formation sur le déroulement des entrevues devrait traiter :

- Des procédures de conduite des entrevues sur le terrain;
- De la supervision des activités de terrain et des procédures d'établissement de rapports;
- De la structure de l'équipe de l'enquête et du rôle de chaque membre de l'équipe.

Évaluation de la formation

45. L'évaluation de la formation doit se faire à plusieurs niveaux. Il faut évaluer les enquêteurs afin de déterminer s'ils sont capables de mener efficacement une entrevue, et de quel appui particulier, le cas échéant, ils ont besoin. Les enquêteurs peuvent eux-même donner leur évaluation de la formation qu'ils ont reçue et de leurs instruc-

teurs. L'évaluation doit être constante et se poursuivre durant toute la période initiale de collecte de données et à la fin du travail sur le terrain.

46. De leur côté, les superviseurs doivent également être évalués par l'équipe centrale de l'enquête. Il faut préciser ici que la nature de la formation doit être adaptée aux tâches que sont censés accomplir les superviseurs, telles que d'essayer de faire revenir un répondant sur un refus, d'effectuer des contre-vérifications de certaines entrevues et d'éviter les réponses. Des protocoles détaillés de ces procédures doivent être rédigés et avoir été clairement expliqués lors de la formation.

47. Les enquêteurs peuvent faire l'objet d'une évaluation formelle à la fin de la formation et une sorte de certificat peut être remis à chaque enquêteur ayant franchi ce cap avec succès. Il appartient à chaque pays de décider lui-même de la procédure qu'il entend suivre à ce sujet.

Tableau X.3

Liste récapitulative à utiliser pour l'examen des procédures de formation

- | |
|---|
| <ul style="list-style-type: none"> • Nombre de sessions de formation • Nombre de jours de formation • Qui a dispensé la formation et quelles étaient ses compétences en formation dans le domaine des enquêtes sur la santé ? • Quelle documentation a-t-elle été utilisée ? • Composantes pratiques : observation des rôles joués dans le contexte réel • Problèmes qui se sont posés lors de la formation • Évaluation de la formation |
|---|

E. Exécution de l'enquête

48. La planification et la gestion d'une enquête sont des tâches complexes, d'un point de vue logistique ou autre, qui exigent beaucoup de préparation, d'organisation et de déploiement de forces sur le terrain pour obtenir l'échantillon désiré. Stratégiquement, l'exécution de l'enquête est un élément clé dont dépend la qualité des données. C'est pourquoi il est très important de veiller avec soin à la qualité d'exécution et au suivi de l'enquête en temps réel, afin de pouvoir traiter les problèmes à mesure qu'ils se posent.

49. L'élément déterminant de la qualité est la façon dont l'enquête est effectivement menée sur le terrain. Cette qualité exige que le pays dispose d'une solide organisation centrale de l'enquête. Il faut que chaque opération (à savoir, l'impression des questionnaires, l'établissement des listes d'échantillonnage, la mobilisation des participants, l'envoi d'équipes d'enquêteurs, l'exécution d'une supervision quotidienne sur le terrain, l'édition des questionnaires, etc.) soit préparée et surveillée avec soin pour en assurer la qualité. En particulier :

- a) Il appartient à chaque équipe de l'enquête d'élaborer un plan d'exécution et un calendrier d'actions décrivant clairement les aspects logistiques de l'enquête. Ce plan doit indiquer le nombre d'enquêteurs nécessaires pour couvrir telle ou telle fraction de l'échantillon d'une région donnée, y compris le nombre de visites (et de secondes visites) et le taux de succès. Autrement dit, il doit tenir compte du taux de non-réponse prévu et des entrevues incomplètes, ainsi que de la présence de l'équipe sur le terrain;

- b) Chaque équipe doit comprendre un superviseur qui supervise et coordonne le travail des enquêteurs et leur dispense sur place formation et appui. Pour l'ESM, le ratio idéal entre superviseurs et enquêteurs varie entre 1:5 et 1:10, selon le pays et selon l'endroit;
- c) Il incombe aux superviseurs de fixer la charge de travail journalière avec les enquêteurs et d'examiner les résultats en fin de journée. Lors de cet examen, les enquêteurs rendent compte à leur superviseur de leurs entrevues de la journée et de leurs résultats. Les superviseurs doivent examiner les entrevues complètes afin de s'assurer que le choix effectué par les enquêteurs des répondants au sein du ménage est correct et que le questionnaire est complet et codé avec précision;
- d) Chaque centre national de direction de l'ESM doit tenir un journal afin de suivre le déroulement de l'enquête. Les éléments à consigner dans ce journal sont les suivants :
 - Le nombre de répondants contactés et les nombres d'entrevues complètes et incomplètes;
 - Les taux de réponse, de refus et de non-contact;
 - Le nombre de secondes visites et les résultats des visites.

Des informations doivent être tenues sur chaque enquêteur, afin que son travail puisse constamment être suivi par le superviseur. Cette base de données sur les enquêteurs peut ensuite servir à tenir ceux-ci informés de leur performance et aider aux futurs recrutements;

- e) Chaque pays doit effectuer une enquête pilote d'une semaine ou deux au début de la période d'enquête. Cette enquête pilote doit être une sorte de répétition générale de l'enquête proprement dite. Cinquante pour cent de l'échantillon pilote serait ensuite réinterrogé par un autre enquêteur afin de démontrer la stabilité d'exécution de l'entrevue. La phase pilote devrait faire l'objet d'un examen critique et d'une discussion avec l'OMS, et ses données devrait être rapidement analysées afin que tout problème d'exécution puisse être identifié. Comme l'instrument à utiliser pour l'enquête aura déjà été soumis à des essais avant cette phase, celle-ci n'aurait pour but que de repérer les problèmes linguistiques et questions de faisabilité d'une importance mineure et de permettre de mieux planifier la phase principale. Elle devrait également permettre d'identifier certaines erreurs particulièrement évidentes dans les questions à omettre. Les résultats de la phase pilote doivent permettre de redresser ces erreurs et de procéder à de légers ajustements. L'étude principale ne devrait commencer qu'après consultation de l'OMS;
- f) Il conviendrait de reconnaître l'utilité de l'impression et du collationnement des questionnaires (par exemple, le code de couleurs utilisé pour les rotations, la lamination des cartes des répondants). Tous les pays devraient envoyer un exemplaire de leurs documents imprimés à l'OMS;
- g) Conformément aux spécifications du contrat de l'ESM, 10 % des répondants devraient être choisis une fois de plus au hasard pour être vérifiés par les superviseurs ou d'autres équipes. Ce contrôle peut se faire par téléphone ou en personne, et doit être structuré de manière que l'entrevue initiale soit conduite convenablement. L'entrevue de contrôle devrait porter sur les informations démographiques de base et sur toute information non recueillie lors de l'entrevue initiale;

- h) Conformément aux spécifications du contrat de l'ESM, une fraction de 10 % de l'échantillon total, choisie au hasard, devrait être soumise une deuxième fois à l'ensemble de l'entrevue par un autre enquêteur dans les sept jours suivant la première entrevue, afin que l'on puisse évaluer la fiabilité du questionnaire (les répondants de la seconde entrevue ne devraient pas être les mêmes que ceux soumis au contrôle spécifié à l'alinéa g ci-dessus;
- i) Les taux de réponse devraient être suivis continuellement et chaque centre devrait recourir à un ensemble de stratégies diverses pour accroître la participation à l'enquête et réduire ces taux. Par exemple, les annonces à la télévision, à la radio, dans les journaux et sur les médias locaux, l'envoi de lettres ou de cartes aux participants, la demande d'assistance aux agents de santé locaux, l'offre d'encouragements à participer, la négociation avec les autorités traditionnelles ou autres responsables locaux, etc., sont autant de techniques de relations publiques pouvant être utilisées pour maximiser le taux de réponse. Le choix des méthodes est laissé aux centres eux-mêmes;
- j) Chaque enquête doit tendre à obtenir le taux de réponse le plus élevé possible. Les spécifications du contrat de l'ESM exigent un taux de réponse global d'au moins 75 %. Ce seuil ne veut pas dire qu'il faille s'arrêter une fois ce seuil atteint. Il indique simplement le minimum acceptable généralement convenu par les collaborateurs à l'ESM au vu des enquêtes passées effectuées dans beaucoup de pays. Souvent, les taux de réponse à l'ESM ont été supérieurs à ce seuil. Le taux de réponse peut varier selon les pays et doit être comparé à celui qu'ont obtenu les mêmes pays pour d'autres enquêtes. Pour le calcul du taux de réponse, il convient d'utiliser la même définition d'une entrevue complète dans tous les pays. Lors des procédures de nettoyage des données appliquées pour déterminer si une entrevue a été complète, on utilise un algorithme fondé sur un ensemble de variables clés;
- k) Rappels : conformément aux spécifications du contrat de l'ESM, les équipes de l'enquête devraient tenter jusqu'à 10 rappels (y compris des appels téléphoniques, des notes ou des cartes indiquant que l'enquêteur est passé). Le nombre moyen de ces rappels dépend du taux de réponse, et chaque centre devrait examiner le gain réalisé par chaque rappel supplémentaire et consulter l'OMS sur le nombre de rappels à juger suffisant dans le pays concerné;
- l) La mise en œuvre de l'enquête dépend en grande partie des ressources disponibles. Chaque enquête doit être évaluée dans le contexte du pays où elle a lieu. Il est essentiel de comparer les résultats avec ceux d'enquêtes comparables dans le même pays. Lors de l'évaluation, il importe de tenir compte des coutumes et traditions locales. Il convient de mesurer les arbitrages entre avoir moins d'enquêteurs faisant plus d'entrevues pendant une période plus longue et avoir plus d'enquêteurs faisant moins d'entrevues pendant une période moins longue.

Tableau X.4
Liste récapitulative des actions entrant dans l'exécution d'une enquête

Enquête pilote

- Où l'enquête pilote a-t-elle été effectuée ?
- Quel type de formation a-t-il été prévu pour l'enquête pilote ?
- Problèmes rencontrés lors de l'entrée des données
- Analyse des données : voir les résultats; quels sont les problèmes qui se sont posés ?
- L'enquête pilote a-t-elle révélé des changements à apporter ?
- Modifications à apporter à la traduction à l'issue de l'enquête pilote

Enquête principale

- Nombre d'enquêteurs, de superviseurs et de coordonnateurs centraux :
 - Comme s'effectue la supervision ? Informations en retour
- Arrangements logistiques :
 - Déplacements : A-t-il été facile de se rendre auprès des ménages ? Quel type de transport a-t-il été utilisé ?
 - Organisation de l'équipe
- Procédures d'établissement de contacts :
 - A-t-il été facile de contacter le répondant ?
 - Combien d'appels ont-ils été nécessaires ?
 - Quel a été le taux de refus et quelle a été la principale raison de refuser l'entrevue ?
- Paiement des enquêteurs
- Signature et enregistrement des formulaires d'acceptation (faisant partie du questionnaire ou sur feuille séparée)
- Procédures de vérification sur le terrain par les superviseurs
- Procédures de vérification à l'échelon central
- Renvoi des questionnaires au bureau central et sécurité
- Vérification finale des questionnaires et des procédures pour corriger les erreurs
- Procédures de vérification et supervision :

Rapports de production hebdomadaires :

- Pour évaluer le processus de l'entrevue
- Pour examiner les taux de réponse, de refus et de non-contact : assurer un taux de réponse
- Pour suivre les résultats et s'assurer de la collecte des données

Vérification des dossiers :

- Le nombre de contacts (et de tentatives de contact) est-il indiqué de façon détaillée ?
 - Est-ce que 10 % au moins des entrevues de chaque enquêteur sont vérifiées par s'assurer que certaines réponses restent constantes (l'âge, l'instruction, la composition du ménage) et que l'entrevue a bien eu lieu ?
 - Vérifier le nombre d'entrevues déjà effectuées et préparer un calendrier d'entrevues
 - Vérifier que les codes indiquant le résultat final pour les entrevues effectuées et les refus ont été attribués correctement
 - S'assurer que les formulaires d'acceptation sont signés

Détacher toutes les données d'identification des questionnaires et du programme d'entrée des données.

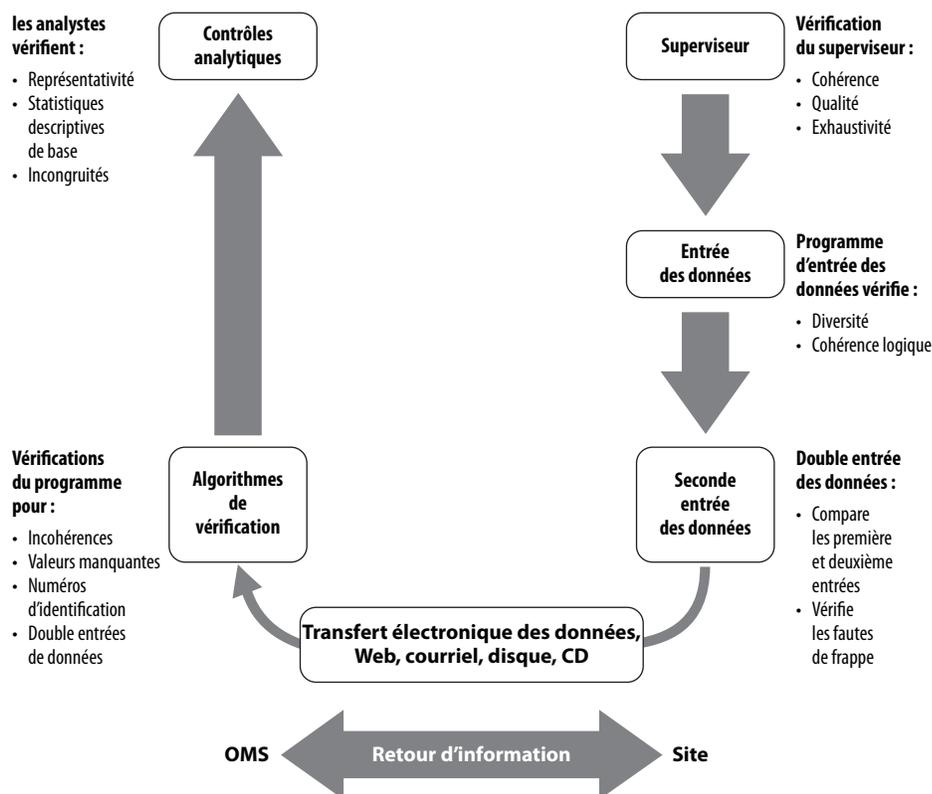
Rédiger un rapport avec des recommandations pour toutes mesures à prendre.

F. Entrée des données

50. Le résultat durable de l'enquête réside dans les données qui sont recueillies. Il importe que ces données soient saisies avec précision et en temps voulu. Le processus d'entrée des données de l'ESM a été conçu de telle manière que les données sont entrées immédiatement à l'échelon local avec coordination à l'échelon central. Il est essentiel que les données soient mises sur ordinateur le plus tôt possible après leur collecte. Ainsi, les vérifications de routine se font aisément à l'aide d'ordinateurs locaux. Toutes erreurs constatées peuvent alors être traitées pendant que l'enquête se poursuit sur le terrain.

51. La figure X.2 ci-dessous décrit le parcours des données de l'ESM et les mesures d'assurance de la qualité correspondantes. Les tâches exécutées à l'OMS sont illustrées dans la partie gauche de cette figure.

Figure X.2
Parcours des données de l'ESM et mesures d'assurance de la qualité correspondantes



52. Une fois que l'entrevue a eu lieu, il est procédé aux opérations suivantes :

- Le superviseur vérifie le formulaire du questionnaire avant l'entrée des données;
- L'entrée des données (ou leur saisie/enregistrement) se fait selon le programme d'entrée de données de l'OMS. Ce programme vérifie les écarts (par exemple, les écarts permis pour les variables de réponse) et s'assure de la cohérence logique des codes correspondants (par exemple, une maladie ne peut durer plus longtemps que l'âge de la personne, et les hommes ne peuvent avoir de problèmes gynécologiques, etc.);

- La seconde entrée de données a pour but d'identifier les fautes de frappe et les questions omises par accident;
- Les données sont envoyées par série à l'OMS par courriel, CD-ROM ou disquette;
- Une fois que les données sont reçues à l'OMS, des programmes vérifient s'il y a des discordances, si certaines valeurs manquent ou si les numéros d'identification posent des problèmes, et vérifient également les cas où il y a des contre-vérifications. Ces programmes permettent d'établir un rapport qui est envoyé aux pays. Par ailleurs, toutes les corrections reçues du pays sont incorporées aux données;
- Des analystes vérifient la représentativité des données, les statistiques descriptives de base et les valeurs qui s'écartent de la norme. Ils vérifient la représentativité en comparant la distribution par âge et par sexe de l'échantillon avec la distribution prévue de la population. Les statistiques descriptives de base sont utilisées pour déterminer la distribution des réponses et pour repérer toute distribution anormale, tout résultat bizarre et toutes valeurs qui s'écartent de la norme;
- L'OMS communique les résultats de ces analyses aux pays. S'il y a lieu, ceux-ci envoient ensuite des corrections et/ou des explications.

53. Questions importantes concernant la qualité de l'entrée des données :

- L'entrée des données doit se faire à l'aide d'un programme conçu à cet effet, qui comporte des contrôles de qualité. L'utilisation de programmes ne comportant pas ces éléments peut être un inconvénient;
- Le superviseur devrait vérifier les formulaires d'entrevue remplis avant que ne commence l'entrée des données;
- Seuls les membres responsables de l'équipe doivent avoir accès au programme d'entrée des données et personne d'autre. C'est essentiel pour préserver la confidentialité des données;
- Les données doivent être entrées deux fois pour éviter les fautes de frappe ou les erreurs. Le programme d'entrée des données indique cette double entrée lorsque la seconde entrée est terminée;
- Les pays doivent bien veiller à entrer le numéro d'identification (ID). Une liste des ID valides est envoyée aux pays. Le programme comprend un contrôle qui permet de s'assurer que le numéro d'identification est entré correctement. L'utilisation de l'ID correct est particulièrement importante pour les cas de contre-vérification, car l'ID est utilisé pour comparer les tests initiaux aux seconds tests;
- Les données doivent être communiquées régulièrement à l'OMS, par exemple, une fois par jour ou une fois par semaine;
- Lorsque l'OMS reçoit les données des pays, elle les vérifie et informe les pays de ses conclusions alors même que se poursuit la collecte des données;
- Certaines règles sont appliquées pour maintenir l'intégrité et la précision des données, qui consistent notamment à déterminer si le même répondant est utilisé deux fois et à mesurer le nombre de données manquantes.

54. Les données d'identification seront détachées des questionnaires et le programme d'entrée des données garde les informations confidentielles dans un fichier séparé. Il incombe au pays de veiller au maintien de la confidentialité. La sécurité des données lors de leur transfert sur l'Internet est assurée par cryptage.

Évaluation de l'entrée des données

55. Il convient de surveiller et de vérifier attentivement les aspects suivants (voir tableau X.5) :

- Le nombre de personnes affectées à l'entrée des données et leur formation;
- Le nombre de formulaires remplis par jour et par personne, y compris les taux d'erreur;
- Les procédures de contrôle et la supervision de l'entrée des données;
- L'intervalle qui s'écoule entre l'entrevue sur le terrain et l'entrée des données;
- Le nombre et la régularité des entrevues envoyées à l'OMS et les problèmes liés à l'envoi des données.

Tableau X.5

Liste récapitulative des vérifications à faire en ce qui concerne le processus d'entrée des données

<ul style="list-style-type: none"> • Quelles sont les personnes qui assurent l'entrée des données ? • Quel est le taux d'achèvement et le taux d'erreur de chacune de ces personnes ? Certaines d'entre elles ont-elles besoin de perfectionnement ? • Observer le processus d'entrée des données. Quel est le système utilisé pour vérifier le nombre de questionnaires affectés à chaque enquêteur ? • Discussion de l'analyse des données et des calculs de la matrice de qualité des données, et nécessité d'un appui supplémentaire • Questionnaires : <p>Choisir plusieurs questionnaires remplis par chaque enquêteur et vérifier que :</p> <ul style="list-style-type: none"> • Les noms ont été retirés des questionnaires • La feuille de couverture a été détachée du questionnaire • Les listes de ménages ont été établies au hasard et remplies comme il convient • L'écriture est lisible et nette • Les options ont été marquées comme il convient (par exemple, les options sont entourées d'un cercle, <i>et non pas</i> cochées, soulignées ou rayées) • S'il y a lieu, il est répondu aux questions ouvertes • Les réponses aux questions ouvertes sont consignées textuellement • Il n'y a pas d'erreurs en ce qui concerne les questions à sauter • Des femmes répondent bien aux questions auxquelles il doit être répondu par des femmes <ul style="list-style-type: none"> • Double entrée des données • Utilisation d'un programme d'entrée des données : <ul style="list-style-type: none"> • Vérifier la confidentialité et la sécurité des données • Les données sont bien entrées deux fois • Vérifier le codage de la base de données par rapport à la copie sur papier • Vérifier les erreurs concernant la portée, la cohérence, l'acheminement et autres • Vérifier le nombre de données manquantes • Le nombre de personnes affectées à l'entrée des données et leur formation • Le nombre de formulaires remplis par jour et par personne, y compris les taux d'erreur • Les procédures de contrôle et la supervision de l'entrée des données • L'intervalle qui s'écoule entre l'entrevue sur le terrain et l'entrée des données • Le nombre et la régularité des entrevues envoyées à l'OMS et les problèmes liés à l'envoi des données
--

56. Bien que l'on puisse atténuer plusieurs problèmes liés à l'entrée des données en procédant par entrevues assistées par ordinateur, méthode qui consiste à entrer les données pendant que se déroule l'entrevue, ces programmes informatiques nécessitent eux-même des vérifications, afin d'assurer le bon déroulement de l'entrevue et la bonne application de toutes les règles concernant les questions à sauter et l'entrée de données se situant dans les fourchettes spécifiées.

G. Analyse des données

57. Avant même une analyse approfondie des données de l'ESM, on dispose d'un certain nombre de vérifications systématiques de leur qualité. L'exécution de ces vérifications, appelée « mesure des données de l'ESM », donne un aperçu des indicateurs de la qualité des données.

58. Les éléments de cette mesure sont :

- L'intégrité, qui comprend le taux de réponse (compte tenu des ménages dont l'admissibilité n'est pas nécessairement connue, auquel cas il faut faire une estimation de la proportion des ménages admissibles ou, si les ménages dont l'admissibilité est douteuse sont exclus du calcul des taux de réponse, il faut donner une claire justification de la supposition selon laquelle ces ménages n'avaient pas de répondants admissibles) et les questionnaires incomplets ainsi que les refus de réponse à certaines questions. On calcule la fréquence des données manquantes au niveau de non-réponse par question parmi tous les répondants et au niveau de répondants refusant de répondre à toutes les questions. Cela permet d'identifier les problèmes d'exécution de l'enquête, en particulier les points du questionnaire qui sont sources de problèmes.
- L'indice de déviation de l'échantillon » [*sampling deviation index* (SDI)], qui mesure le degré selon lequel l'échantillon s'écarte de la représentativité de la population cible. Si cet indice est élevé, l'analyse doit être stratifiée. On peut évaluer le SDI en utilisant la distribution chi-carrée. Si certains groupes clés ont été intentionnellement suréchantillonnés, il faut en tenir compte afin d'ajuster le SDI en lui appliquant le coefficient de suréchantillonnage désiré.
- La fiabilité, qui indique la possibilité de reproduire les résultats en appliquant le même instrument de mesure au même répondant à différents stades et en faisant appel à différents enquêteurs. Cette analyse utilise les données du protocole d'essai/répétition d'essai appliqué dans 50 % des entrevues pilotes et dans 10 % de l'ensemble de l'échantillon.
- La comparaison avec des indicateurs de validité externes, c'est-à-dire avec d'autres résultats d'autres enquêtes, telles que le recensement, ainsi qu'avec des données des secteurs public et privé.

59. Ces mesures sont examinées de façon plus approfondie dans la section suivante. Le traitement des données se fait au niveau du pays, où se trouve la capacité nécessaire, ainsi qu'au siège de l'OMS.

60. Une autre analyse des données au niveau du pays est considérée comme un moyen essentiel d'assurer l'utilisation efficace des résultats. Le siège et les bureaux régionaux de l'OMS identifient les pays qui ont besoin d'appui pour l'analyse complète des données et élaborent des mécanismes pour dispenser cet appui.

Évaluation de l'analyse des données

61. L'évaluation de cet aspect nécessite une vérification de l'existence dans le pays des compétences nécessaires pour effectuer cette analyse et du niveau d'appui dont le pays peut avoir besoin ou que le pays peut apporter à d'autres pays.

H. Indicateurs de qualité

62. Il est utile de mesurer le degré d'assurance de qualité au moyen d'indicateurs. Ces indicateurs peuvent ensuite être utilisés pour mesurer d'autres facteurs liés au contexte qui influent sur la qualité, de manière à compléter le cycle d'évaluation de la qualité. À notre connaissance, il n'a pas été proposé d'ensemble systématique d'in-

dicateurs pour suivre la qualité d'une enquête et en rendre compte. L'ESM utilise certains indicateurs chiffrables expliqués ci-dessous ainsi qu'une évaluation qualitative structurée selon le processus d'examen par des pairs pour rendre compte de l'assurance de qualité.

63. En règle générale, toute enquête sur les ménages est sujette à deux types d'erreurs : l'erreur d'échantillonnage et l'erreur autre que l'erreur d'échantillonnage. Il y a erreur d'échantillonnage lorsqu'une enquête est effectuée sur un échantillon de la population plutôt que sur la population tout entière. Elle est fonction de la taille de l'échantillon, de la variabilité qui caractérise la population pour les grandeurs mesurées et d'autres aspects de la conception de l'échantillon tels que la stratification et la mise en grappes. Les erreurs autres que les erreurs d'échantillonnage, en revanche, sont fonction de facteurs tels que la nature des concepts étudiés, la précision et le degré d'intégrité du cadre d'échantillonnage, la fidélité des procédures de sélection sur le terrain vis-à-vis de la conception de l'échantillon et les erreurs d'exécution de l'enquête. Ce dernier facteur est lié à certains problèmes tels que la mauvaise conception du questionnaire, les erreurs commises par les enquêteurs lorsqu'ils posent les questions et les erreurs commises par les répondants lorsqu'ils y répondent, l'entrée des données et d'autres erreurs de traitement, la non-réponse et des techniques d'estimation inexactes. Certaines des erreurs autres que les erreurs d'échantillonnage qui peuvent se mesurer et se chiffrer sont indiquées ci-dessous.

64. En ce qui concerne le suivi du résultat final de l'enquête, les indicateurs types suivants sont actuellement utilisés pour mesurer la qualité des données de l'ESM.

1. Indice de déviation de l'échantillon

65. L'indice de déviation de l'échantillon (SDI)³ indique les proportions de l'échantillon par âge et par sexe par rapport aux données sur la population émanant d'une source indépendante, ces dernières étant considérées comme la norme. L'ESM a utilisé comme source indépendante la base de données des Nations Unies sur la population, mais on peut utiliser toute autre source plus récente et fiable de données sur la population. Le SDI est un indicateur de la qualité des données de l'échantillon, c'est-à-dire de leur représentativité de l'ensemble de la population. Un ratio de 1 signifie que l'échantillon de l'enquête coïncide avec les caractéristiques de la population tout entière en ce qui concerne la répartition par âge et par sexe, tandis que toute déviation par rapport à l'unité indique un suréchantillonnage ou un sous-échantillonnage d'une tranche d'âge ou d'un sexe.

66. La valeur attendue de 1 (représentativité idéale) est rarement observée dans les enquêtes, en raison des erreurs d'échantillonnage. La figure X.3 illustre le SDI de l'une des enquêtes, qui révèle une sous-représentation des tranches d'âge jeunes et une sur-représentation des personnes plus âgées, surtout des hommes.

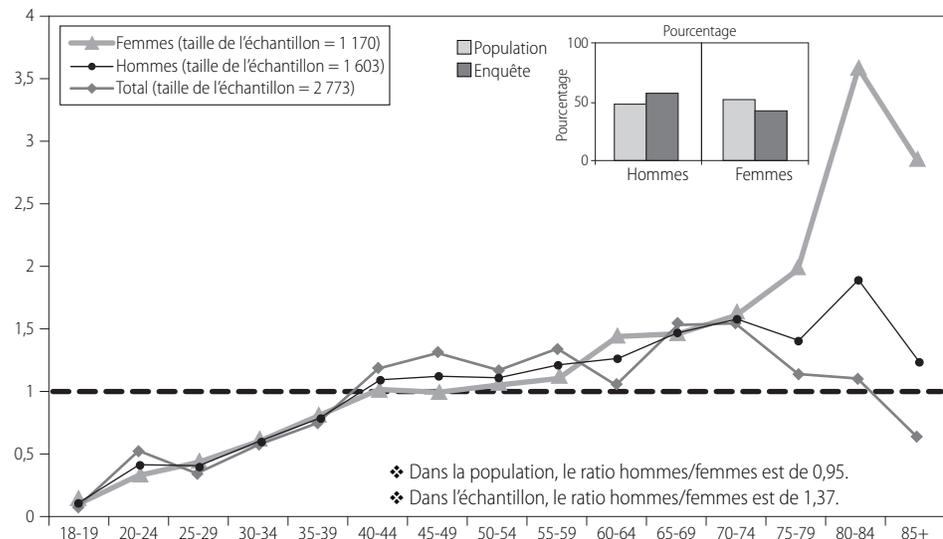
2. Taux de réponse

67. Le taux de réponse indique le taux d'achèvement des entrevues de l'échantillon sélectionné, c'est-à-dire le nombre d'entrevues complètes avec des personnes ou des ménages admissibles (par exemple, la sélection d'un logement pour le « ménage » censé l'habiter qui se révèle n'être autre qu'un logement vacant n'est pas admissible). Cet indicateur indique la qualité de l'enquête, la qualité idéale étant un taux de réponse de 100 %. Un taux de réponse de 60 % est généralement considéré comme le minimum acceptable, et l'ESM exige un taux de réponse d'au moins 75 %.

$$SDI = \sum_{a=1}^a |1 - indice_a|$$

où a = tranches d'âge et où l'indice est le ratio de l'échantillon de la tranche d'âge à la population de cette tranche d'âge telle qu'elle figure dans la base de données des Nations Unies ou d'autres sources actualisées telles que le recensement de la population du pays. Cet indice mesure la représentativité de l'échantillon des points de vue de la distribution par âge et par sexe. La cohérence de cet indice peut être vérifiée par les tests chi-carré ou pi-star.

Figure X.3.
Exemple d'indice de déviation d'un échantillon



3. Pourcentage de données manquantes

68. Le pourcentage de données manquantes se définit comme la proportion de réponses manquantes à la suite d'une entrevue avec un répondant. L'ESM mesure la proportion de personnes ne donnant pas le minimum de réponses acceptable (par exemple, 10 % dans les entrevues face à face) pour déterminer la qualité des entrevues. Ce pourcentage permet également d'identifier les questions à problème, qui donnent lieu à un fort pourcentage de non-réponse (plus de 5 %) parmi les répondants admissibles.

4. Taux de concordance entre la première et la seconde entrevue

69. Les taux de concordance entre la première et la seconde entrevue indiquent la stabilité mesurée à la variabilité des réponses données à deux reprises. On les calcule comme taux de concordance corrigé des effets du hasard (c'est-à-dire comme statistique kappa de coefficients de corrélation par catégorie pour des variables continues). Cet indicateur mesure à quel point une question donne les mêmes réponses lorsqu'elle est posée lors de deux entrevues. En général, un score supérieur à 0,4 est jugé acceptable; un score supérieur à 0,6 est jugé satisfaisant; et un score supérieur à 0,8 est jugé excellent (6).

70. Pour chaque statistique clé de l'enquête, le principal indicateur de la qualité, mesuré à l'erreur observée dans les données de l'échantillon, est l'erreur estimative type. Celle-ci donne une estimation de la fourchette dans laquelle se situe l'erreur d'échantillonnage (par exemple, $\pm 3\%$) autour d'une estimation donnée. On calcule également, si possible, le coefficient d'effet de la conception pour les échantillons en grappes à plusieurs degrés de l'ESM. Ce coefficient est le ratio de la variance entre l'échantillon effectif et un échantillon aléatoire simple hypothétique de même taille. Étant donné qu'un véritable échantillon aléatoire n'est pas pratique dans les enquêtes à grande échelle, pour des questions de coût (notamment de coûts de transport), on calcule généralement une variance d'échantillonnage (carré de l'erreur type) à titre de comparaison avec un échantillon aléatoire (Kish, 1995b). Un effet de conception

compris entre 1 et 6 est généralement jugé acceptable pour les indicateurs jugés intéressants pour l'ESM.

I. Les rapports des pays

71. L'une des dimensions importantes de l'assurance de qualité a trait au résultat final, c'est-à-dire à la notification des données, en raison de l'impact de l'enquête mesurée à ce qu'elle ajoute à notre base de connaissances et à son influence sur la politique à venir. Bien entendu, la qualité du rapport est liée à l'utilité de l'ESM pour les besoins du pays. Les résultats de l'ESM sont présentés dans un certain nombre de rapports différents, à savoir :

- a) Les rapports établis par chaque pays participant à l'ESM :
 - i) Les comptes rendus analytiques à l'intention des décideurs et du public;
 - ii) Le rapport détaillé à l'intention des chercheurs et autres utilisateurs scientifiques;
- b) Les rapports régionaux et internationaux sur certaines questions spécifiques.

72. Le canevas initial pour le rapport de pays [71, a, ci-dessus] comprend :

- Une introduction (contenant notamment les informations devant servir à définir la politique et les informations disponibles sur les systèmes de santé);
- Une analyse de l'exécution de l'enquête (comprenant, par exemple, une description de l'enquête, des méthodes d'échantillonnage, de la formation, de la collecte et du traitement des données, des procédures d'assurance de qualité, de la métrologie de l'enquête);
- Un aperçu des résultats de l'enquête et de ses conséquences pour la politique (portant, par exemple, sur les éléments à prendre en compte pour le système de santé, les caractéristiques de la population et des ménages, la couverture des interventions en matière de santé, la santé de la population, la faculté de réponse des systèmes de santé, les dépenses de santé);
- Des conclusions : des recommandations spécifiques pour la politique de santé et le suivi des Objectifs du Millénaire pour le développement dans le pays.

73. Ce canevas sera ensuite développé en collaboration étroite avec les pays, les bureaux régionaux et d'autres parties intéressées.

74. Il convient de définir clairement une stratégie de diffusion du rapport à travers les médias, des ateliers et d'autres manifestations. Il faut engager les différentes parties prenantes à utiliser les informations recueillies dans le cadre de l'enquête à travers des débats de fond.

75. Les pays eux-mêmes doivent être les principaux responsables de la rédaction de ces rapports. L'OMS les aide en fournissant les données et l'appui technique essentiels aux équipes de rédaction ainsi que les outils nécessaires à la préparation et à la discussion de ces rapports.

76. L'ESM fournit des informations sur divers aspects de la santé de la population et des systèmes de santé. Ces informations portent sur de nombreux éléments du cadre d'évaluation de la performance de ces systèmes. Par ailleurs, ces enquêtes fournissent des informations détaillées sur d'autres aspects, tels que certains facteurs de risque, les fonctions des systèmes de santé, l'épidémiologie de certaines maladies et les services de santé. Il est donc important de recueillir les meilleures informations possibles des données de l'ESM.

77. Certains pays peuvent également chercher à utiliser les données de l'ESM pour des analyses sous-régionales. Le plus souvent, cela nécessite de plus grands échantillons. Parfois, les données de l'ESM peuvent être utilisées avec celles d'autres sources telles que le recensement et d'autres enquêtes.

78. À long terme, on prévoit que la structure modulaire de l'ESM permettra d'intégrer diverses enquêtes sur la santé et les systèmes de santé en une même enquête.

Évaluation des rapports des pays

79. L'analyse des données et la rédaction des rapports par les pays est l'aboutissement de l'enquête. La qualité des rapports et la façon dont les résultats sont analysés déterminent la manière dont les enquêtes à venir seront exécutées ainsi que l'impact de leurs résultats sur l'élaboration et le suivi de la politique à l'intérieur du pays.

J. Visites sur place

80. Les pays de l'ESM savent d'avance ce que l'on attend d'eux pour l'exécution de l'enquête et l'assurance de sa qualité. À cet égard, il importe de documenter le travail de terrain. Pour ce faire, l'OMS engagera des conseillers indépendants en assurance de qualité qui effectueront des visites dans chaque pays. Ces visites sur place constitueront en fait un examen externe par des pairs de l'exécution de l'enquête et un moyen de contrôler le respect des normes d'assurance de qualité. Ces visites sur place seront également l'occasion de signaler les problèmes et de les résoudre dès le début du processus. L'équipe pays et le conseiller en assurance de qualité établiront ainsi une évaluation structurée de la qualité d'ensemble de l'enquête selon les directives de l'OMS.

81. L'assurance de qualité est un processus qui ne peut se réduire à une simple visite sur place. La relation entre les conseillers en assurance de qualité et les équipes pays doit être perçue comme un processus à long terme en trois phases : avant, pendant et après la visite sur place.

82. Avant la visite sur place, le pays et le conseiller en assurance de qualité doivent préparer un dossier pour la visite conforme aux directives fondamentales de l'OMS en matière d'assurance de qualité, portant sur tous les aspects indiqués sur la liste récapitulative des actions prévues pour cette visite. Ce dossier doit contenir toutes les informations générales disponibles sur le site, l'organisme chargé de l'enquête, la conception de l'échantillon, les compétences locales, les instruments et la formation utilisés sur place et le canevas pour le rapport du pays sur l'ESM. Les informations non disponibles devront être recueillies lors de la visite.

83. Les responsables pays au siège de l'OMS et les conseillers en assurance de qualité doivent être en communication directe avec l'enquêteur principal ou le principal responsable de l'enquête dans le pays, afin que le processus d'assurance de qualité fasse partie intégrante du processus d'exécution de l'enquête. Cela permet de donner aux enquêtes une culture d'assurance de qualité. Le processus d'assurance de qualité a pour but non pas de contrôler ou de policer l'ESM mais d'assurer sa qualité en lui dispensant une assistance et un appui.

84. Pour que la visite sur place ait le maximum d'impact, il faut qu'elle soit prévue vers la fin de la formation et le début de la collecte de données. Cette visite doit porter sur tous les aspects du processus d'enquête; autrement dit, elle doit diagnosti-

quer les problèmes, leur suggérer des remèdes, tenir compte du contexte local, apporter un appui et établir une relation permanente.

85. Les conseillers en assurance de qualité [*quality assurance advisers* (QAAs)] lorsqu'ils visitent les pays, ont pour fonction de diagnostiquer les problèmes et de souligner les points forts dans l'exécution de l'enquête. Ils ont principalement pour tâche d'examiner le processus d'exécution de l'ESM utilisé dans le pays et de signaler tout écart par rapport aux normes prévues d'assurance de qualité. Leur conclusion sur le point de savoir si l'écart est important et comment il pourrait y être remédié est essentielle. Le QAA doit également apporter un appui direct par ses entretiens avec le siège de l'OMS ou veiller à ce que cet appui soit dispensé par une autre entité.

86. Les QAA effectuent leur évaluation selon une liste récapitulative structurée comprenant les diverses opérations à effectuer, par ordre d'importance. Cette évaluation doit comprendre l'analyse de la « métrologie de l'enquête » (dans la mesure où certaines données sont déjà entrées au moment où a lieu leur visite), c'est-à-dire des indicateurs de la qualité des données.

87. L'évaluation de l'assurance de la qualité est débattue conjointement avec l'équipe chargée de l'enquête et l'OMS. Les pays doivent savoir à l'avance ce que l'on attend d'eux en ce qui concerne les procédures d'assurance de la qualité.

88. Le rapport sur la visite sur place est suivi du rapport du pays sur l'ESM, qui est le produit final de la visite sur place et de l'appui au pays. La visite sur place entame le processus de rédaction du rapport du pays et examine certaines stratégies pour sa production, y compris la façon d'utiliser les conclusions de l'enquête pour la définition d'une politique.

K. Conclusions

89. L'assurance de qualité est l'une des questions clés de l'exécution d'une enquête. Il est nécessaire et possible de spécifier certains mécanismes d'assurance de la qualité pour chaque phase de l'enquête. Si ces mécanismes sont définis concrètement, on peut mesurer la qualité de chacune de ces phases et suivre celle de l'enquête dans son ensemble.

90. L'établissement de l'assurance de qualité nécessite un changement dans l'état d'esprit des personnes chargées de l'enquête, car l'examen et l'évaluation sont obligatoires à chacun des stades de son exécution.

91. Il est essentiel d'évaluer constamment la qualité des indicateurs durant toute l'enquête. Ce processus ne doit pas être considéré simplement comme a posteriori; il doit être utilisé pour apporter des corrections à mi-parcours, au fur et à mesure que des problèmes sont décelés. Cette importante amélioration continue de la qualité doit être intégrée à toutes les enquêtes.

92. L'informatique permet aujourd'hui d'élaborer un système de gestion et de suivi de l'enquête tout au long de son exécution, et de renforcer la confiance inspirée par ses données.

93. Il importe de documenter systématiquement les questions critiques (par exemple, les questions concernant l'exécution de l'enquête, la formation, etc.) par des rapports qualitatifs et des indicateurs quantitatifs (tels que l'indice de déviation de l'échantillon, les taux de réponse, les pourcentages de données manquantes et la concordance entre la première et la seconde entrevue), afin de donner aux usagers des données des renseignements essentiels sur la qualité de l'enquête.

94. Le but recherché par le processus d'assurance de qualité est une enquête qui produise des données de meilleure qualité, qui puissent être présentées comme valides, fiables et comparables.

95. L'application continue de ces procédures d'assurance de qualité permet de fixer des normes de qualité pour la collecte de données à l'échelle internationale, et les méthodes de suivi de ces normes sont appelées à évoluer continuellement.

REMERCIEMENTS

Nous tenons à exprimer notre vive gratitude aux experts ci-après, de divers pays et institutions, pour leur contribution à l'établissement de directives d'assurance de qualité pour l'ESM :

D^r Farid Abolhassani, République islamique d'Iran
 D^r Sergio Aguilar-Gaxiola, États-Unis d'Amérique
 D^r Atalay Alem, Éthiopie
 D^r Lorna Bailie, Canada
 D^r Russell Blamey, Australie
 D^r Carlos Gomez-Restrepo, Colombie
 D^r Oye Gureje, Nigéria
 D^r Holub Jiri, République tchèque
 M. Mark Isserow, Afrique du Sud
 D^r Feng Jiang, Chine
 M. Jean-Louis Lanoe, France
 P^r Howard Meltzer, Royaume-Uni de Grande-Bretagne et d'Irlande du Nord
 M. Steve Motlatla, Afrique du Sud
 M^{me} Lipika Nanda, Inde
 D^r Kültegin Ögel, Turquie
 D^r Gustavo Olaiz Fernandez, Mexique
 D^r Mhamed Ouakrim, Maroc
 D^r Jorun Ramm, Norvège
 D^r Wafa Salloum, République arabe syrienne
 D^r Shen Mingming, Chine
 D^r Benjamin Vicente, Chili

Consultants en échantillonnage

P^r Steve Heeringa, Université du Michigan, Institute of Social Research, États-Unis d'Amérique
 P^r Nanjamma Chinnappa, Inde, ex-président de l'International Association of Survey Statisticians

Conseillers régionaux de l'OMS

M^{me} M. Mohale M., Conseillère régionale auprès du Bureau régional de l'OMS pour l'Afrique
 D^r Siddiqi Sameen, Conseiller régional auprès du Bureau régional de l'OMS pour la Méditerranée orientale
 D^r Amina Elghamry, Conseiller régional auprès du Bureau régional de l'OMS pour la Méditerranée orientale
 D^r Lars Moller, Conseiller régional auprès du Bureau régional de l'OMS pour l'Europe

D^r Myint Htwe, Conseiller régional auprès du Bureau régional de l'OMS pour l'Asie du Sud-Est

Dr. Soe Nyunt-U, Conseiller régional auprès du Bureau régional de l'OMS pour le Pacifique occidental

RÉFÉRENCES

- Biemer, P. P., *et al.*, eds. (1991). *Measurement Errors in Surveys*. New York, Wiley.
- Bryant, B. E. (1975). Respondent selection in a time of changing household composition. *Journal of Marketing Research*, vol. 12, pp. 129-135.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, vol. 20, pp. 37-46.
- DeLepper, M. H., H. Scholten et R. Stern, eds (1995). *The Added Value of Geographical Information Systems in Public and Environmental Health* Dordrecht, Netherlands: Kluwer Academic Publishers.
- Fleiss, J. L. (1981). *Statistical Methods for Rates and Proportions*, 2^e éd. New York, John Wiley and Sons.
- Kish, L. (1995a). *Survey Sampling*. New York, John Wiley and Sons.
- _____ (1995b) Methods for design effects. *Journal of Official Statistics*, vol. 11, pp. 55-77.
- Lyberg, L. E., *et al.*, eds. (1997). *Survey Measurement and Process Quality*. New York, Wiley.
- Organisation mondiale de la santé (2000). *Rapport sur la santé dans le monde*. Genève, OMS.
- _____ (2002). *World Health Survey: Quality Assurance and Guidelines: Procedures for Quality Assurance Implementation by Country Survey Teams and Quality Assurance Advise*. Genève, OMS.
- Statistique Canada (1998). *Quality Guidelines*, 3^e éd. Ottawa.
- Üstün, T. B. *et al.* (2001). *Disability and Culture; Universalism and Diversity*. Göttingen, Allemagne, Hogrefe Huber.
- _____ (2003a). WHO Multi-country Survey Study on Health and Responsiveness 2000-2001. Dans *Health System Performance Assessment: Debates, Methods and Empiricism* (C. J. L. Murray et D.B. Evans, eds.). Genève, OMS.
- _____ (2003b). The World Health Surveys. Dans *Health System Performance Assessment: Debates, Methods and Empiricism* (C. J. L. Murray et D. B. Evans, eds.). Genève : OMS.
- _____ (2003c). *World Health Organization Disability Assessment Schedule II (WHO DAS II): Development and Psychometric Testing*. Genève, OMS. In collaboration with WHO/National Institute of Health Joint Project Collaborators.
- Valentine, N. B., A. de Silva et C. J. L. Murray (2000). *Estimating Responsiveness Level and Distribution for 191 Countries: Methods and Results*. Global Programme on Evidence Discussion Paper Series, No. 22. Genève, OMS.

Chapitre XI

Établissement de rapports et compensation des erreurs autres que les erreurs d'échantillonnage au Brésil : pratique actuelle et défis pour l'avenir

PEDRO LUIS DO NASCIMENTO SILVA

Escola Nacional de Ciências Estadísticas/
Instituto Brasileiro de Geografia e Estatística
(ENCE/IBGE)
Rio de Janeiro, Brésil

RÉSUMÉ

Le présent chapitre offre un examen de certaines pratiques actuelles en matière d'établissement de rapports et de compensation des erreurs autres que les erreurs d'échantillonnage au Brésil, et considère trois catégories d'erreurs : les erreurs de couverture, les erreurs dues à la non-réponse et les erreurs de mesure et de traitement. Il présente également certains facteurs qui ajoutent à la difficulté d'accorder une plus grande attention à la mesure et à la maîtrise des erreurs autres que les erreurs d'échantillonnage. Par ailleurs, il évoque certaines initiatives récentes qui devraient permettre d'améliorer la situation.

Termes clés : processus d'enquête, couverture, erreurs de mesure, établissement de rapports d'enquête, qualité des données.

A. Introduction

1. Il convient de définir la notion d'erreur telle qu'elle s'applique à une statistique ou à une estimation d'une grandeur ou d'un paramètre cible inconnu. Il s'agit de la différence entre l'estimation (disons, \hat{y}) et la « vraie valeur théorique » (disons, Y) d'un paramètre qui serait obtenue ou signalée si toutes les sources d'erreur étaient éliminées. Peut-être, comme certains le prétendent, serait-il préférable de parler de déviation [voir discussion dans Platek et Särndal (2001, sect. 5)]. Cependant, le terme erreur est si ancré dans les esprits que nous n'essaierons pas de l'éviter. Ici, nous voulons parler des erreurs des enquêtes, c'est-à-dire des erreurs d'estimation fondées sur les données d'enquêtes. D'après Lyberg *et al.* (1997, p. xiii), « les erreurs dans les enquêtes se divisent en deux grandes catégories : les erreurs d'échantillonnage et les erreurs autres que les erreurs d'échantillonnage ». Dans la terminologie moderne, l'analyse des erreurs dans les enquêtes s'inscrit dans le débat plus général sur la qualité des données.

2. Pour illustrer ce concept, supposons que l'estimation du revenu mensuel moyen d'une certaine population, tel qu'il ressort d'une enquête, est de 900 dollars des États-Unis et que le revenu mensuel moyen effectif des divers membres de cette population, obtenu à l'issue d'une énumération complète sans erreurs de notification et de traitement, est de 850 dollars des États-Unis. Dans cet exemple, l'erreur de l'estimation serait de + 50 dollars des États-Unis. En général, on ne remarque pas les erreurs dans les enquêtes car les valeurs vraies des paramètres ne se remarquent pas (ou ne peuvent se remarquer). L'un des cas au moins où les erreurs d'échantillonnage des estimations statistiques peuvent se remarquer est celui de l'échantillonnage à partir de fichiers d'ordinateur, où l'on peut calculer, si besoin est, les différences entre les estimations et les valeurs calculées à l'aide des ensembles complets de données. Les échantillons de fichiers tirés d'un recensement de la population destinés à l'usage du public offrent un exemple d'application pratique. Au Brésil, des échantillons de ce type sont sélectionnés à partir des données des recensements depuis 1970. Néanmoins, ce type de situation est l'exception plutôt que la règle.

3. Les erreurs d'échantillonnage se définissent comme la différence entre les estimations fondées sur une enquête par sondage et les grandeurs correspondantes de la population qui seraient obtenues si un recensement était effectué selon les mêmes méthodes de mesure; elles « sont causées par l'observation d'un échantillon au lieu de la population tout entière » (Särndal, Swensson et Wretman, 1992, p. 16). « Les erreurs autres que les erreurs d'échantillonnage sont toutes les autres erreurs » (ibid) survenant dans une enquête. Elles peuvent se produire et se produisent effectivement dans toutes sortes d'enquêtes, y compris les recensements. Dans les recensements et les enquêtes utilisant de vastes échantillons, les erreurs autres que les erreurs d'échantillonnage sont la principale source d'erreurs dont on doit se préoccuper.

4. Les estimations des enquêtes sont sujettes à deux types d'erreurs : les distorsions et les erreurs liées aux variables. La distorsion est une erreur commise sur la valeur attendue de l'estimation de l'enquête, qui écarte cette valeur de la valeur vraie du paramètre visé. Les erreurs liées aux variables sont les écarts qui se produisent sur les estimations de l'enquête lorsque le processus est répété. En ce qui concerne les erreurs d'échantillonnage, on évite généralement la distorsion ou on la rend négligeable en utilisant des procédures d'échantillonnage, des tailles d'échantillons et des méthodes d'estimation adéquates. Ainsi, la marge est le principal aspect de la distribution de l'erreur d'échantillonnage qu'il faut prendre en compte. Le paramètre clé pour décrire cette marge est l'erreur type, c'est-à-dire l'écart type dans la distribution de l'erreur d'échantillonnage.

5. Les erreurs autres que les erreurs d'échantillonnage se classent en deux grandes catégories (Särndal, Swensson et Wretman, 1992, p. 16) : « les erreurs dues à la non-observation » et « les erreurs d'observation ». Les erreurs dues à la non-observation résultent du défaut d'obtention des données requises auprès de certaines parties de la population cible (erreurs de couverture) ou d'une partie de l'échantillon sélectionné (erreurs de non-réponse). Les erreurs de couverture ou du cadre désignent les inclusions, omissions et répétitions indues de certaines unités dans le cadre de l'enquête, qui entraînent un excès ou une insuffisance de couverture de la population cible. Les erreurs de non-réponse sont causées par le défaut d'obtention de données pour certaines unités sélectionnées pour l'enquête. Les erreurs d'observation peuvent être de trois types : erreurs de spécification, erreurs de mesure et erreurs de traitement. Biemer et Fecso (1995, chap. 15) définissent les erreurs de spécification comme celles qui se produisent lorsque : « 1) les concepts analysés ne sont pas mesurables ou sont mal définis; 2) les objectifs de l'enquête sont insuffisamment spécifiés; ou 3) les données recueillies ne correspondent pas aux concepts spécifiés ou aux variables ci-

bles ». Les erreurs de mesure correspondent aux valeurs observées pour les questions et variables de l'enquête après la collecte des données qui s'écartent des valeurs vraies correspondantes qui seraient obtenues si des méthodes de mesure idéales étaient utilisées. Les erreurs de traitement sont celles qui sont introduites lors du traitement des données recueillies, c'est-à-dire lors du codage, de la frappe, de la pondération et de la tabulation des données de l'enquête. Tous ces types d'erreurs sont traités dans les sous-sections de la section B, à l'exception des erreurs de spécification. L'exclusion de ces dernières ne veut pas dire qu'elles ne sont pas importantes, mais seulement que la discussion et le traitement de ces erreurs ne sont pas bien établis au Brésil.

6. D'autres approches de la classification des erreurs autres que les erreurs d'échantillonnage sont présentées dans un manuel des Nations Unies (voir Nations Unies, 1982). Dans certains cas, il n'y a pas de claire distinction entre les erreurs de non-réponse, de couverture et de mesure; c'est ce qui se produit dans une enquête par sondage à plusieurs degrés sur les ménages lorsque l'on manque un membre d'un ménage recensé. S'agit-il d'une erreur de mesure, de non-réponse ou de couverture ?

7. Les erreurs autres que les erreurs d'échantillonnage peuvent aussi se diviser en variances et distorsions non dues à l'échantillonnage. La variance non due à l'échantillonnage mesure la variation des estimations d'une enquête si l'on soumet le même échantillon à des répétitions hypothétiques du processus d'enquête dans des conditions essentiellement semblables (Nations Unies, 1982, p. 20). La distorsion non due à l'échantillonnage correspond aux erreurs qui résultent du processus d'enquête et des conditions de l'enquête et débouchent sur des estimations dont la valeur attendue diffère de la valeur vraie du paramètre. Comme exemple de la distorsion non due à l'échantillonnage, supposons que des membres d'une population tendent à indiquer un revenu inférieur à leur revenu réel de 30 % en moyenne. Dans ce cas, indépendamment de la conception de l'échantillon et des procédures d'estimation, en l'absence de toute information extérieure, les estimations du revenu moyen données par l'enquête seraient en moyenne de 30 % inférieures à la valeur vraie du revenu moyen de ces membres de la population. La majeure partie du présent chapitre traite des mesures tendant à éviter ou à compenser la distorsion non due à l'échantillonnage.

8. Les questions de qualité des données retiennent davantage l'attention depuis quelques années, et ont donné lieu à un certain nombre d'initiatives et de publications, ainsi qu'à des conférences internationales (voir section D). Malheureusement, le débat se limite essentiellement aux pays développés, avec très peu de participation et de contribution des pays en développement et des pays en transition. Telle est la principale conclusion à laquelle aboutit l'examen des comptes rendus et des publications de ces diverses conférences et initiatives. Néanmoins, plusieurs études ont été publiées récemment sur ce sujet à propos d'enquêtes effectuées dans les pays en transition dans le journal *Statistics in Transition* (Kordos, 2002), mais ce journal ne semble guère répandu dans les bibliothèques des pays en développement.

9. En ce qui concerne les erreurs d'échantillonnage, il existe une théorie unifiée de la mesure et de l'estimation [voir, par exemple, Särndal, Swensson et Wretman (1992)], appuyée par la large diffusion des méthodes et techniques d'échantillonnage probabiliste comme norme pour l'échantillonnage dans les enquêtes (Kalton, 2002), et par la généralisation d'un logiciel type qui permet l'application pratique de cette théorie aux enquêtes réelles. Si les échantillons sont convenablement prélevés et collectés, les estimations de la variabilité d'échantillonnage d'une enquête sont relativement faciles à établir. C'est déjà ce qui se fait pour de nombreuses enquêtes effectuées dans les

pays en développement et les pays en transition, encore que cette pratique soit loin de devenir une norme imposée.

10. Quoi qu'il en soit, la diffusion et l'analyse de ces mesures de la variabilité ont encore beaucoup de chemin à parcourir. Dans nombre d'enquêtes, les estimations des erreurs d'échantillonnage ne sont ni calculées ni publiées, ou ne le sont que pour un petit nombre de variables ou d'estimations. En général, elles ne le sont pas pour la plupart des estimations de l'enquête car cela implique un énorme travail de calcul. Si cela rend difficile à tout utilisateur extérieur l'évaluation du degré de variabilité de l'échantillonnage d'une variable donnée, il est néanmoins possible de se faire une idée de son ordre de grandeur en comparant ce degré avec celui d'une variable similaire dont on a estimé l'erreur type. Souvent, les commentaires sur les estimations d'une enquête ne font pas cas du degré de variabilité de ces estimations. Par exemple, l'Enquête mensuelle sur la main-d'œuvre brésilienne (Instituto Brasileiro de Geografia e Estatística, 2002b), lancée en 1980, calcule et publie chaque mois des estimations des coefficients de variations (CV) des indicateurs avancés dont l'enquête établit l'estimation. Toutefois, il n'est pas établi d'estimations des erreurs types pour les différences affichées par ces indicateurs entre deux mois de suite ou pour le même mois examiné deux années de suite. Or, la plupart des commentaires publiés chaque mois avec les estimations portent sur les différences (les variations des indicateurs mensuels). Ces estimations des erreurs types sur les estimations des variations ne sont calculées que depuis peu pour l'analyse interne [voir Correa, Silva et Freitas (2002)], et elles ne sont pas encore mises à la disposition des utilisateurs extérieurs des résultats des enquêtes. Tel est aussi le cas des estimations « complexes », telles que les indicateurs du marché du travail corrigées des variations saisonnières.

11. Si la situation est loin d'être idéale pour les erreurs d'échantillonnage, alors que la théorie et le logiciel sont amplement disponibles et que la culture de l'échantillonnage est largement répandue, le traitement des erreurs autres que les erreurs d'échantillonnage dans les enquêtes sur les ménages et autres enquêtes effectuées dans les pays en développement est beaucoup moins développé. L'absence d'une théorie largement acceptée sur les erreurs autres que les erreurs d'échantillonnage [voir Lyberg et al. (1997, p. xiii); Platek et Särndal (2001), et les débats ultérieurs], l'absence de méthodes types pour la collecte d'informations sur les éléments constitutifs de ces erreurs et l'estimation de leurs paramètres, et l'absence d'une culture qui reconnaisse l'importance de la mesure, de l'évaluation et de la notification de ces erreurs font qu'elles retiennent moins l'attention dans les enquêtes effectuées dans les pays en développement ou les pays en transition. Cela ne veut pas dire que la plupart des enquêtes effectuées dans les pays en développement ou les pays en transition sont de mauvaise qualité, mais plutôt que l'on sait peu de choses de leurs niveaux de qualité.

12. Compte tenu de ce qui précède sur l'état de la mesure de l'erreur autre que l'erreur d'échantillonnage dans les enquêtes effectuées dans les pays en développement et les pays en transition, nous examinerons à présent l'état de la pratique actuelle (section B) dans l'expérience brésilienne. Bien que cet examen se limite à un seul pays (le Brésil), nous estimons qu'il présente un intérêt pour les statisticiens d'autres pays en développement, compte tenu du fait que les écrits sur ce sujet sont peu nombreux. Nous indiquerons ensuite les défis à venir pour l'amélioration de la pratique des enquêtes dans les pays en développement et les pays en transition (section C), toujours en nous fondant sur la pratique de ces enquêtes au Brésil.

B. Pratique actuelle de la notification et de la compensation des erreurs autres que les erreurs d'échantillonnage dans les enquêtes sur les ménages au Brésil

13. Au Brésil, les principales grandes enquêtes par sondage sur les ménages sont effectuées par l'Instituto Brasileiro de Geografia e Estatística (IBGE), l'institut central de statistique du Brésil. Pour aider le lecteur à comprendre les références à ces enquêtes nous en présentons les principales caractéristiques, la couverture et la périodicité au tableau XI.1 ci-dessous.

Tableau XI.1
Quelques caractéristiques des principales enquêtes par sondage sur les ménages au Brésil

Nom de l'enquête	Périodicité	Population couverte	Thème/sujet
Recensement de la population	Tous les 10 ans (le dernier en 2000)	Habitants de logements privés et collectifs dans le pays	Caractéristiques du ménage, situation de famille, fécondité, mortalité, religion, race, éducation, travail, revenu
Enquête nationale par sondage sur les ménages [National Household Sample Survey (PNAD)]	Annuelle, sauf les années de recensement	Habitants de logements privés et collectifs dans le pays, sauf dans les zones rurales du Nord	Caractéristiques du ménage, religion, race, éducation, travail, revenu et suppléments spéciaux sur divers sujets
Enquête mensuelle sur la main-d'œuvre	Mensuelle	Habitants de logements privés de six grandes zones métropolitaines	Éducation, travail, revenu
Enquête sur les dépenses des ménages (POF)	1974-1975, 1986-1987, 1995-1996, 2002-2003	Nationale pour l'édition 2002-2003; 11 grandes zones métropolitaines lors des deux précédentes éditions; nationales pour l'édition 1974-1975	Caractéristiques du ménage, dépenses et revenus des ménages
Enquête sur les niveaux de vie (PPV)	1996-1997	Habitants de logements privés des régions Nord-Est et Sud-Est	Vaste couverture de questions liées à la mesure du niveau de vie
Enquête sur l'économie informelle en milieu urbain (ECINF)	1997	Membres de ménages privés participant à l'économie informelle dans les zones urbaines	Travail, revenu et caractéristiques du travail effectué dans l'économie informelle

1. Erreurs de couverture

14. Les erreurs de couverture se traduisent par une insuffisance ou un excès de couverture des unités de population. Il y a insuffisance de couverture lorsque des unités de la population cible ne figurent pas dans le cadre et sont donc inaccessibles pour l'enquête. Inversement, il y a excès de couverture quand des unités n'appartenant pas à la population cible figurent dans le cadre et qu'il n'y a aucun moyen de les distinguer des unités admissibles avant l'échantillonnage, ainsi que quand le cadre inclut deux fois des unités admissibles. Les erreurs de couverture résultent également d'une classification erronée des unités dans les strates par suite d'informations inexactes ou périmées (par exemple, lorsqu'un logement est exclu du processus d'échantillonnage pour non-occupation alors qu'en fait il était occupé au moment de l'enquête). L'insuffisance de couverture est généralement plus préjudiciable que l'excès de couverture pour les estimations de l'enquête. Il n'y a aucun moyen de recouvrer les unités manquantes, alors que les unités improprement comptées peuvent souvent être identifiées lors du travail de terrain ou du traitement des données et les corrections et ajustements appropriés peuvent être effectués; cela dit, les unités improprement incluses alourdissent le coût de l'enquête par unité admissible.

15. Les problèmes de couverture sont souvent jugés plus graves pour le recensement que pour les enquêtes par sondage, car pour les recensements, on n'a pas à se soucier des erreurs d'échantillonnage. Toutefois, c'est là une idée fautive. Dans certai-

nes enquêtes par sondage, la couverture peut parfois poser un problème aussi grave, sinon plus, que l'erreur d'échantillonnage. Par exemple, les enquêtes par sondage excluent parfois du processus d'échantillonnage des unités de certaines régions difficiles d'accès ou de catégories difficiles à sonder (et leur confèrent donc une probabilité d'inclusion nulle). Ces exclusions peuvent s'expliquer par le souci de préserver la sécurité de l'enquêteur (notamment dans les zones de conflit ou de violence) ou de limiter le coût (par exemple, lorsqu'il faut se rendre dans des régions dont l'accès est d'un coût prohibitif ou prendrait trop de temps). Si la définition de la population cible n'indique pas clairement ces exclusions, l'enquête risque de souffrir d'une insuffisance de couverture. Ces problèmes se traduisent à leur tour par une distorsion des estimations, car les unités exclues de la population étudiée tendent à être différentes des unités incluses. Lorsque l'enquête cherche à couvrir ces populations difficilement accessibles, il faut une planification spéciale pour s'assurer que la couverture s'étend à ces groupes de la population cible ou de la population pour laquelle on cherche à tirer des conclusions.

16. Il se pose un problème du même ordre lorsque l'on effectue des enquêtes répétées dans des pays à la couverture téléphonique insuffisante, ou à taux d'analphabétisme élevé, où la collecte de données doit se faire par entrevues face à face. Lorsque des enquêtes disposent d'une courte période pour les entrevues, leur couverture doit souvent se limiter aux zones faciles d'accès. Au Brésil, par exemple, l'Enquête mensuelle sur la main-d'œuvre (PME) ne s'effectue que dans six grandes zones métropolitaines (Instituto Brasileiro de Geografia e Estatística, 2002b). Sa définition limitée de la population cible est l'une des principales sources de critique de la validité de cette enquête : avec une population cible qui est trop restreinte pour de nombreux usages, cette enquête ne donne pas d'indications de l'évolution de l'emploi et du chômage dans les autres régions du pays. Bien qu'elle rende compte correctement de la situation de la « population cible » vivant dans ces six zones métropolitaines, de nombreux utilisateurs interprètent à tort ses chiffres comme s'ils se rapportaient à l'ensemble de la population du Brésil. Le pays prévoit de modifier cette enquête pour répondre à ce problème dans son édition 2003-2004. Des problèmes similaires se posent pour d'autres enquêtes, par exemple les Enquêtes brésiliennes sur les revenus et les dépenses de 1987-1988 et 1995-1996 (dont la couverture était limitée à 11 zones métropolitaines) et l'Enquête sur la mesure des niveaux de vie (LSMS) de 1996-1997 (dont la couverture était limitée aux régions Nord-Est et Sud-Est seulement). Dans une moindre mesure, tel a été également le cas de la grande enquête « nationale » par sondage sur les ménages effectuée au Brésil (Instituto Brasileiro de Geografia e Estatística, 2002a). Cette enquête ne couvre pas les zones rurales du nord du pays, en raison de leurs coûts d'accès prohibitifs. Bianchini et Albieri (1998) donnent une analyse plus détaillée de la méthodologie et de la couverture de diverses enquêtes sur les ménages effectuées au Brésil.

17. Des problèmes similaires se posent pour de nombreuses enquêtes effectuées dans d'autres pays en développement ou en transition, où la couverture fréquente de certaines zones difficiles d'accès risque d'être trop coûteuse. La règle importante à suivre à cet égard est d'inclure dans toute publication fondée sur une enquête une indication précise de la population effectivement couverte, suivie d'une description éventuelle des sous-groupes potentiellement importants qui en ont été exclus.

18. Les mesures de l'erreur de couverture ne sont pas toujours publiées avec les estimations de l'enquête, ce qui permettrait aux utilisateurs extérieurs d'effectuer une évaluation impartiale de l'impact des problèmes de couverture dans leur analyse. Ces mesures ne sont généralement données que lorsque les chiffres des recensements de population sont publiés, à peu près tous les dix ans et, même dans ce cas, elles ne sont

pas directement associées au problème de couverture des enquêtes sur les ménages effectuées au cours de la décennie précédente.

19. Au Brésil, la seule « enquête » où une analyse plus complète de la couverture est effectuée est le recensement de la population. Celle-ci consiste généralement en un ensemble d'enquêtes par sondage après énumération et d'analyses démographiques. Une enquête par sondage après énumération (PES) est une enquête effectuée principalement pour évaluer la couverture du recensement ou d'une enquête similaire, mais dans beaucoup de pays, la PES est souvent utilisée pour évaluer également le contexte de l'enquête. Au Brésil, la PES qui a suivi le recensement de la population de l'année 2000 a porté sur quelque 1 000 zones d'énumération qu'elle a sondées en faisant appel à une équipe distincte et indépendante d'énumérateurs qui ont dû suivre les mêmes procédures que celles qui avaient été suivies par les énumérateurs qui ont effectué le recensement. Une fois recueillies les données de la PES, on les compare aux données obtenues pour les unités correspondantes lors du recensement. On utilise ensuite les résultats de cette comparaison pour une double estimation [voir, par exemple, Marks (1973)], qui produit des estimations de l'insuffisance de couverture telles que celles indiquées au tableau XI.2 ci-après. L'analyse démographique de la population et des mouvements de population fondée sur les registres administratifs des naissances et des décès permet également de vérifier les comptages de la population du recensement et d'évaluer leur degré de couverture. Au Brésil, cette pratique n'est valable que dans certains États du Sud et du Sud-Est, où les registres des naissances et des décès sont suffisamment précis pour donner des informations utiles à cet égard.

20. Leur coût élevé est l'un des principaux obstacles à l'application généralisée des enquêtes PES pour l'estimation de la couverture et pour l'analyse du recensement. Ces enquêtes doivent être préparées et exécutées avec soin pour produire des résultats fiables. Par ailleurs, il faut que leurs résultats soient plus ou moins fractionnés, faute de quoi elles ne sont que d'une utilité limitée. Dans certains cas, les ressources nécessaires à une telle enquête ne peuvent être obtenues, et dans d'autres, les responsables de la planification du recensement peuvent penser que ces ressources seraient utilisées plus efficacement si elles étaient affectées à l'amélioration du recensement lui-même. Or, il est difficile, sinon impossible, d'améliorer le recensement sans évaluer les principaux problèmes et détecter où ils se situent. La PES permet de localiser les principales sources des problèmes de couverture et d'indiquer les aspects de la collecte des données qui ont besoin d'être améliorés pour les recensements futurs, ainsi que d'estimer l'ampleur de l'insuffisance de couverture afin de la compenser. C'est pourquoi nous recommandons vivement que, lors de l'établissement du budget et de la préparation du recensement, des ressources soient réservées pour permettre d'effectuer une PES d'ampleur raisonnable juste après la collecte de données pour le recensement. L'évaluation de la couverture d'une analyse démographique est généralement moins coûteuse qu'une PES, mais elle nécessite à la fois un accès à des sources de données extérieures et une connaissance des méthodes employées pour cette analyse. Toutefois, dans la mesure du possible, on devrait prévoir un budget pour ce type d'étude et du temps pour l'effectuer dans le cadre de l'évaluation du recensement.

21. Dans la plupart des pays, développés ou non, on n'ajuste pas les chiffres des recensements pour tenir compte de l'insuffisance de couverture, car il n'existe pas de théorie ou de méthode largement acceptées pour corriger les erreurs de couverture, et les estimations de l'insuffisance de couverture des PES n'est pas suffisamment fiable, ou parce que certains facteurs politiques empêchent de modifier les estimations du recensement ou bien pour une combinaison de ces facteurs ou d'autres encore. De ce fait, les estimations de la population publiées à partir des données du recensement n'offrent aucune compensation pour l'insuffisance de couverture. Dans certains cas,

les informations sur cette insuffisance, si elles existent, doivent être traitées comme « confidentielles » et ne peuvent être accessibles au public, car la perception de ce type d'informations pourrait nuire à la crédibilité des résultats du recensement si elles sont mal interprétées. Nous recommandons d'éviter cette pratique et de publier les résultats de la PES ou de les tenir à la disposition des utilisateurs du recensement.

22. L'analyse ci-dessus se rapporte aux enquêtes à large couverture de la population. Le problème d'une bonne évaluation de la couverture est encore plus délicat pour certains sous-groupes de population tels que les minorités ethniques ou autres, car la taille de l'échantillon nécessaire à une PES exige généralement des ressources qui débordent le budget disponible. On sait très peu de choses de la qualité de la couverture de ces catégories de population par les recensements et autres enquêtes sur les ménages dans les pays en développement. Au Brésil, aucune enquête effectuée après recensement depuis le recensement de 1970 n'a donné d'estimation utilisable des groupes ethniques ou autres minorités pertinentes. Ces estimations se sont situées en dessous des chiffres réels, tant pour les ménages que pour les individus, et ont été fractionnées par grandes zones géographiques (États). Les résultats des estimations de cette insuffisance de couverture lors du recensement de 2000 ont été publiés récemment (Oliveira *et al.*, 2003). Nous ne présentons ici que les résultats à l'échelle du pays, y compris les estimations des taux d'omission de ménages et de personnes lors des recensements de 1991 et 2000. Les taux d'insuffisance de couverture étaient voisins pour ces deux recensements, avec des taux légèrement inférieurs pour le recensement de 2000. L'une des recommandations pour améliorer la PES après les recensements de la population brésilienne a été d'accroître l'estimation de l'insuffisance de couverture de manière à inclure certaines sous-catégories de population pertinentes, tels que les groupes ethniques ou les groupes d'âge.

23. Les chiffres du tableau XI.2 sont plus élevés que ceux indiqués pour des recensements similaires dans certains pays développés. Les taux d'omission révèlent un degré d'insuffisance de couverture non négligeable. Jusqu'ici, les résultats des recensements de la population brésilienne sont publiés, comme dans la plupart des pays, sans ajustements pour tenir compte de cette insuffisance. Toutefois, de tels ajustements sont effectués par la suite, en fonction des projections de population publiées après le recensement. Il faut que la recherche évalue l'impact possible de l'ajustement des estimations des recensements pour tenir compte de l'insuffisance de couverture, et il faut que cette recherche s'accompagne de discussions, de planification et de décisions quant à la fiabilité à attendre des estimations des PES pour qu'elles puissent être utilisées à cette fin.

Tableau XI.2

Estimation des taux d'omission dans les recensements de la population du Brésil découlant des enquêtes après les recensements de 1991 et 2000 (en pourcentage)

Couverture par catégorie	Recensement de 1991	Recensement de 2000
Logements privés occupés	4,5	4,4
Personnes vivant dans des logements privés non omis	4,0	2,6
Personnes omises	8,3	7,9

Source : Oliveira *et al.* (2003).

2. Non-réponse

24. L'expression « non-réponse » est utilisée pour désigner les cas où des données sont absentes pour certaines unités (non-réponse unitaire), pour certaines unités après plusieurs entrevues (non-réponse répétée) ou pour certaines variables parmi

plusieurs unités (non-réponse ponctuelle). La non-réponse touche toutes les enquêtes, qu'il s'agisse d'un recensement ou d'un sondage. Elle peut également influencer sur les données émanant de sources administratives qui sont utilisées pour des études statistiques. La plupart des enquêtes utilisent certaines procédures opérationnelles pour éviter ou réduire l'incidence de la non-réponse. Celle-ci est plus gênante lorsque la réponse n'est pas « aléatoire » (différences de taux de non-réponse entre d'importants sous-groupes de population et lorsque les taux de réponse sont faibles. Si la non-réponse est aléatoire, elle a pour principal effet d'accroître la variance des estimations de l'enquête car la taille de l'échantillon se trouve réduite. Toutefois, si la participation à une enquête (réponse) est fonction de certains traits et caractéristiques des répondants et/ou des enquêteurs, le problème le plus préoccupant devient la distorsion, surtout si les taux de non-réponse sont élevés.

25. Selon Särndal, Swensson et Wretman (1992, p. 575) : « Les principales techniques pour faire face à la non-réponse sont l'ajustement des pondérations et l'imputation. Les ajustements des pondérations se traduisent par un accroissement des pondérations appliquées à l'estimation des valeurs- y des répondants pour compenser les pertes de valeur dues à la non-réponse. L'imputation implique la substitution de « bonnes » valeurs artificielles aux valeurs manquantes. »

26. Des trois types de non-réponse, la non-réponse unitaire est la plus difficile à compenser, car les cadres d'échantillonnage et les archives offrent généralement très peu d'informations à utiliser à cette fin. La méthode de compensation la plus couramment utilisée pour contrer les effets négatifs de la non-réponse unitaire est l'ajustement des pondérations, qui consiste à accroître les pondérations des unités répondantes pour compenser la perte d'unités due à la non-réponse, mais même ce type très simple de compensation n'est pas toujours appliqué. La compensation des cas de non-réponse répétée et de non-réponse ponctuelle se fait souvent par imputation, car dans ces cas, les unités non répondantes ont fourni certaines informations qui peuvent être utilisées pour guider l'imputation et réduire ainsi la distorsion (voir Kalton, 1983, 1986).

27. La non-réponse a plusieurs causes. Elle peut être due au fait que certaines unités n'ont pas été contactées faute de temps, parce qu'elles étaient difficiles à recenser ou parce que les répondants n'étaient pas chez eux. Elle peut aussi être le résultat d'un refus de coopérer ou de l'incapacité à répondre ou à participer à l'enquête. La non-réponse due au refus est souvent faible dans les enquêtes sur les ménages effectuées dans les pays en développement, essentiellement parce que, les citoyens ayant moins d'instruction, ils sont moins enclins à refuser de coopérer aux enquêtes, et parce que les taux d'analphabétisme plus élevés obligent à recueillir les données par des entrevues face à face plutôt que par téléphone ou par envoi de questionnaires par la poste. Ces deux facteurs contribuent à réduire les taux de refus de coopérer, et peuvent aussi déboucher sur des différences de non-réponse selon le niveau d'instruction et de revenu, les éléments les plus instruits et aux revenus plus élevés étant plus enclins à refuser de répondre. En même temps, la réponse ou la participation aux enquêtes ne se traduit pas nécessairement par une plus grande précision dans les réponses : souvent, les taux de réponse plus élevés peuvent en fait masquer les fausses déclarations en réponse à certaines questions, notamment celles qui concernent les variables relatives aux revenus ou à la richesse, par méfiance à l'égard des pouvoirs publics.

28. Les recensements effectués dans les pays en développement souffrent des effets de la non-réponse. Le Brésil utilise deux types de questionnaires pour les recensements : un questionnaire abrégé, qui ne comprend que quelques questions d'ordre démographique (sexe, âge, lien de parenté avec le chef de famille et alphabétisme), et

un questionnaire plus détaillé, comprenant des questions d'ordre socio-économique (race, religion, instruction, travail, revenu, fécondité, mortalité, etc.), en plus des questions figurant sur le questionnaire abrégé. Le questionnaire détaillé est utilisé pour les ménages choisis par un échantillon probabiliste de ménages dans chaque zone d'énumération. Le taux d'échantillonnage est plus élevé (1 sur 5) dans les petites municipalités et plus faible (1 sur 10) dans les municipalités dont la population est estimée à 15 000 habitants ou plus l'année du recensement. Le taux global de non-réponse est très faible (d'environ 0,8 % lors du recensement national de 2000). Toutefois, pour les variables du questionnaire abrégé (qui nécessitent une réponse de tous les ménages participants), il n'est pas prévu de compensation pour la non-réponse, et ce, pour trois raisons : premièrement, la non-réponse est jugée très faible; deuxièmement, on dispose de trop peu d'informations sur les ménages qui ne répondent pas pour que la compensation soit efficace; et troisièmement, il n'y a pas de cadre naturel pour effectuer des ajustements de pondération dans un recensement. De plus, la solution de l'imputation, qui consiste à substituer de « bonnes » valeurs artificielles aux valeurs manquantes, n'est guère prisée, pour les deux premières raisons mentionnées, mais aussi en raison du préjugé défavorable au recours à l'imputation dans de tels cas. Pour les estimations tirées de l'échantillon, on utilise des ajustements de pondération fondés sur des méthodes d'étalonnage pour compenser en partie les cas de non-réponse unitaire.

29. Une approche similaire a été adoptée pour certaines enquêtes sur les ménages. Au Brésil, deux des principales enquêtes sur les ménages, l'Enquête nationale annuelle par sondage sur les ménages (PNAD) et l'Enquête mensuelle sur la main-d'œuvre (PME), utilisent des méthodes non spécifiques de compensation des cas de non-réponse (voir Bianchini et Albieri, 1998). Les seuls ajustements apportés aux pondérations des unités répondantes se font par étalonnage appliqué à l'ensemble de la population de la zone métropolitaine ou de l'État, et ne peuvent donc compenser les différences de non-réponse au sein de certains groupes de population définis par le sexe ou l'âge, par exemple. Les raisons de cette impossibilité sont liées principalement à des considérations opérationnelles, telles que le maintien d'un logiciel particulier utilisé pour les estimations qui a été conçu il y a longtemps, et le fait qu'il est plus simple de ne pas tenir compte des cas de non-réponse. Ces deux enquêtes enregistrent leurs niveaux de non-réponse, mais les informations sur ce point ne sont pas diffusées dans les publications présentant les principaux résultats de ces enquêtes. Néanmoins, des fichiers de microdonnées sont établis pour permettre d'établir des estimations des taux de non-réponse, car les cas des unités qui ne répondent pas sont également inclus dans ces fichiers avec des codes appropriés qui indiquent les raisons de la non-réponse. La PME a fait l'objet d'une récente refonte (Instituto Brasileiro de Geografia e Estatística, 2002b) et a commencé à utiliser au moins une simple méthode de repondération pour compenser le taux de non-réponse unitaire observé. Parmi les nouvelles modifications envisagées pourrait figurer l'introduction d'estimateurs d'étalonnage qui tenteraient de corriger les taux différentiels de non-réponse selon l'âge et le sexe. Toutefois, les études pertinentes, motivées par la constatation que la non-réponse est l'une des causes probables des rotations de la distorsion selon les groupes (Pfeffermann, Silva et Freitas, 2000) dans les estimations mensuelles du taux de chômage, n'en sont encore qu'à un stade préliminaire.

30. L'Enquête sur les dépenses des ménages brésiliens (POF) [dont le dernier cycle remonte à 1995-1996, le nouveau cycle de 2002-2003 étant en cours] a appliqué des méthodes perfectionnées d'ajustement pour non-réponse. Cette enquête utilise un ensemble de méthodes de repondération et d'imputation pour compenser les cas de non-réponse (Bianchini et Albieri, 1998). Les ajustements des pondérations per-

mettent de compenser la non-réponse unitaire, tandis que les méthodes d'imputation permettent de compenser les variables ou blocs de variables pour lesquels les réponses font défaut après collecte et traitement des données. L'attention accrue portée au traitement de la non-réponse s'explique par les taux plus élevés de non-réponse observés pour cette enquête par comparaison avec les enquêtes générales sur les ménages. Un taux de non-réponse plus élevé est à prévoir, compte tenu du fardeau de réponses beaucoup plus lourd imposé par ce type d'enquête (les ménages doivent recevoir au moins deux visites et sont invités à tenir une comptabilité détaillée de leurs dépenses pendant une période de deux semaines). Les rapports sur la méthodologie des enquêtes doivent contenir une analyse des cas de non-réponse, mais pas les publications présentant leurs principaux résultats.

31. Une autre enquête effectuée au Brésil, l'Enquête sur la mesure des niveaux de vie (PPV), qui faisait partie du Programme d'enquêtes sur la mesure des niveaux de vie de la Banque mondiale, avait recours à la substitution de ménages pour compenser les cas de non-réponse. Au Brésil, cette pratique est rarement utilisée, et aucune grande enquête sur les ménages ne l'a adoptée.

32. Après examen de ces diverses enquêtes effectuées à l'intérieur d'un même pays, on constate qu'il n'y a pas d'approche type pour compenser la non-réponse unitaire et en rendre compte. Les méthodes de traitement de la non-réponse varient selon les enquêtes, en fonction des niveaux de non-réponse observés et de l'adhésion de l'enquête à des recommandations internationales, ainsi que du besoin perçu et de la capacité de mise en œuvre de méthodes et procédures de compensation. L'une des méthodes qui pourrait être utilisée pour améliorer cette situation serait de préparer régulièrement des rapports sur les « profils de qualité » des enquêtes sur les ménages. Cette méthode pourrait se révéler souvent plus pratique et plus utile que celle qui consisterait à tenter d'inclure toutes les informations dont on dispose sur les méthodes employées et les limitations des données du recensement ou des publications sur les enquêtes.

33. La situation est très différente dans les cas de non-réponse ponctuelle. Dans les recensements de la population brésilienne, à partir de 1980, on a utilisé les méthodes d'imputation pour combler les lacunes et remplacer les valeurs anormales décelées par application des règles d'édition spécifiées par les spécialistes. En 1991 et 2000, le Brésil a utilisé une combinaison des méthodes d'imputation et des méthodes de Fellegi-Holt, appliquées dans un logiciel tel que DIA (*Detección e Imputación Automática de datos*) [García Rubio et Criado, 1990] et NIM (*New Imputation Methodology*) [Poirier, Bankier et Lachance, 2001], pour effectuer une édition et une imputation intégrées des formulaires de recensement abrégés et détaillés. En 2000, l'imputation a porté non seulement sur les variables catégoriques, mais aussi sur les variables du revenu, grâce à l'utilisation des méthodes des arbres de régression pour trouver les informations qui ont pu être utilisées pour remplacer les informations manquantes sur les revenus. Ce recensement de la population brésilienne a été le premier qui ne comportait aucune valeur manquante dans les fichiers de microdonnées à la fin du traitement des données. La stratégie d'édition et d'imputation appliquée au recensement de la population est bien documentée, quoique la plupart des informations sur le niveau d'édition et d'imputation ne figurent que dans des rapports spécialisés. Pour faciliter l'accès à ces rapports, il est recommandé de les diffuser sur l'Internet.

34. Le traitement des données manquantes ou douteuses dans les autres enquêtes sur les ménages n'est pas aussi élaboré. Pour la PNAD et la PME, on utilise des programmes informatiques pour la détection des erreurs, mais il faut encore une forte dose « d'édition manuelle » et les méthodes d'imputation assistée par ordinateur sont

peu utilisées pour compenser les cas de non-réponse ponctuelle. Si certaines informations font défaut à la fin de la phase d'édition, elles sont codées comme « inconnues ». Les progrès réalisés ces dernières années ont porté sur l'intégration des opérations d'édition à l'entrée des données, ce qui a permis de réduire le coût et le temps de traitement. L'apparition d'ordinateurs portables moins coûteux et plus performants a permis à l'IBGE de progresser encore davantage sur la voie de l'intégration. La version révisée de la PME pour la décennie des années 2000 a permis d'entreprendre en octobre 2001 la collecte d'un échantillon parallèle de la même taille que celui utilisé par l'enquête proprement dite, pour lequel les données ont été obtenues lors d'entrevues face à face assistées par ordinateur (*palmtop*). Il n'y a pas encore de rapport définitif sur la performance des ordinateurs *palmtop*, mais après quelques mois, il semble que la collecte des données se déroulait sans problèmes. Cette technologie a permis aux responsables des enquêtes de se concentrer sur l'amélioration de la qualité à la source, en intégrant toutes les instructions sur les questions à sauter et les contrôles de validité à l'instrument de collecte des données, de manière à éviter d'entrer des fautes de frappe et autres erreurs dans la source. On compense la non-réponse aux questions sur le revenu en utilisant les méthodes des arbres de régression pour trouver des réponses, comme dans le recensement de la population. Toutefois, les résultats de cette nouvelle enquête ne sont connus que depuis peu, et la collecte des données s'est faite parallèlement à l'ancienne série pendant toute une année avant que les données aient été diffusées et que la nouvelle série remplace l'ancienne. L'évaluation plus large et plus détaillée des résultats de cette nouvelle approche de la collecte et du traitement des données se poursuit encore.

35. Dans la PME, chaque ménage est maintenu dans l'échantillon pendant deux périodes de quatre mois chacune, espacées de huit mois. Ainsi, en principe, les données de la première série d'entrevues complètes peuvent être utilisées pour compenser les cas de non-réponse répétée chaque fois qu'un ménage ou un membre d'un ménage n'a pu être contacté à aucun des cycles d'entrevues après le premier. Cette utilisation des données ne se produit pas dans l'ancienne série, pas plus qu'elle n'est prévue pour la nouvelle, et pourtant, elle représente un progrès que les responsables d'enquêtes feraient bien de considérer.

36. La tendance qui se dégage d'une analyse de l'édition et de l'imputation pratiquées dans le cadre de plusieurs enquêtes pour la non-réponse ponctuelle et les données incohérentes ou douteuses est une tendance à la non-standardisation, les méthodes employées variant selon les enquêtes. Les recensements ont clairement été l'occasion d'amples applications de méthodes d'édition et d'imputation automatiques, ce qui n'a pas été souvent le cas des enquêtes de plus petite envergure. Peut-être la taille de l'enquête y est-elle pour quelque chose, en ce sens que l'investissement dans la conception et l'application de méthodes et de procédures acceptables d'imputation automatique est justifiable pour les recensements, mais pas pour des enquêtes de portée plus restreinte, qui disposent en outre de moins de temps pour produire leurs résultats. Pour une enquête répétée comme la PME au Brésil, malgré le temps très court dont elle dispose pour produire ses résultats, il y aurait probablement à gagner à investir davantage dans des méthodes d'édition et d'imputation des données, en raison des perspectives d'exploitation de cet investissement sur plusieurs cycles d'enquête successifs.

3. Erreurs de mesure et de traitement

37. Les erreurs de mesure et de traitement se manifestent par des valeurs obtenues en réponse à des questions ou pour des variables qui, après collecte et traitement

des données, différent des valeurs vraies correspondantes qui aurait été obtenues par l'utilisation de méthodes de mesure et de traitement idéales.

38. Cette question est probablement l'une de celles qui retiennent le moins l'attention des points de vue de la mesure, de la compensation et de la notification dans les enquêtes sur les ménages effectuées dans les pays en développement et les pays en transition. Plusieurs faits récents peuvent être perçus comme pointant vers une amélioration des pratiques de l'enquête et vers une réduction des erreurs de mesure. Premièrement, l'utilisation de méthodes de collecte de données assistée par ordinateur a permis de réduire les erreurs de transcription, dans la mesure où les réponses sont directement entrées dans l'ordinateur et sont immédiatement disponibles pour l'édition et l'analyse. En outre, le flux des questions est contrôlé par l'ordinateur et peut être rendu tributaire des réponses, de manière à éviter les erreurs introduites par l'enquêteur. Les réponses peuvent être examinées en regard des fourchettes attendues, voire des précédentes réponses du même répondant. Les données douteuses ou surprenantes peuvent être signalées et il peut être demandé à l'enquêteur d'interroger le répondant à leur sujet. Ainsi, en principe, les données doivent être de meilleure qualité et moins sujettes à erreur de mesure. Toutefois, on a peu d'indications d'une amélioration de la qualité entre les entrevues assistées par ordinateur et les questionnaires à remplir sur papier, si ce n'est l'avantage de réduire les taux de questions manquées et le nombre de grandeurs se situant hors des fourchettes normales.

39. Un autre progrès provient de la conception et de l'application généralisée d'un logiciel pour l'édition et l'imputation des données (Criado et Cabria, 1990). Comme on l'a déjà indiqué à la section B, les recensements de population ont adopté un logiciel d'édition et d'imputation automatiques pour détecter et compenser les erreurs de mesure et certains types d'erreurs de traitement (par exemple, les erreurs de codage et de frappe) et, en même temps, les cas de non-réponse ponctuelle. Ce logiciel est aussi appliqué pour certaines enquêtes par sondage. Toutefois, le type de compensation appliqué ne permet de détecter que les erreurs « fortuites ». Les erreurs systématiques sont rarement détectées ou compensées par un logiciel d'édition de type courant.

40. La conception de logiciels de codage assisté par ordinateur, ainsi que de logiciels et de matériels de saisie des données, pourrait également contribuer à réduire les erreurs de traitement dans les enquêtes.

41. Si certains progrès ont été enregistrés dans la prévention des erreurs de mesure et de traitement, il n'en est pas de même de l'application des méthodes de mesure, de compensation et de notification des erreurs de mesure. En ce qui concerne les erreurs de mesure, on se concentre principalement sur la prévention, et une fois que le nécessaire a été fait à cet égard, on ne s'attarde guère à mesurer le succès de la préparation et de l'exécution de l'enquête. L'absence de théorie universelle de la mesure complique l'établissement d'objectifs de qualité et l'évaluation de la réalisation de ces objectifs. Par exemple, bien que nous ayons eu des plans d'échantillonnage pour la définition de l'échantillon et des coefficients de variation (erreur type relative) de certaines estimations clés inférieurs à une valeur spécifiée d'avance, nous voyons rarement des plans de collecte et de traitement des données qui visent à maintenir l'imputation en dessous d'un niveau donné, ou qui visent à maintenir certaines mesures dans une marge de tolérance spécifiée (c'est-à-dire une déviation maximum) par rapport à des « valeurs vraies » à haute probabilité. Il peut s'avérer difficile de pouvoir fixer d'avance des objectifs quantitatifs pour tous les types d'erreurs autres que les erreurs d'échantillonnage; cependant, nous estimons que les organismes d'enquête devraient

au moins s'efforcer de mesurer ces erreurs et d'utiliser ces mesures pour fixer des objectifs d'amélioration à venir et de suivre la progression vers ces objectifs.

C. Défis et perspectives

42. Après plus de 50 ans de popularisation des enquêtes par sondage en tant qu'instrument d'observation dans le domaine des sciences sociales, la notion d'erreurs d'échantillonnage et leur maîtrise, leur mesure et leur interprétation ont atteint un certain niveau de maturité, bien que, comme nous l'avons noté, les résultats de nombreuses enquêtes à travers le monde soient publiés sans estimation de ces erreurs. Les progrès ont été beaucoup plus limités en ce qui concerne les erreurs autres que les erreurs d'échantillonnage, du moins pour les enquêtes effectuées dans les pays en développement. Cela n'est pas le seul fait du hasard. Le problème posé par ces erreurs est difficile. Tout d'abord, ces erreurs proviennent de nombreuses sources. Les méthodes de prévention sont tributaires non seulement de la technologie, mais aussi de la culture et du milieu, de sorte qu'il est très difficile de généraliser et de propager les résultats d'expériences fructueuses. Les méthodes de compensation sont généralement complexes et coûteuses à appliquer convenablement. Ces erreurs sont difficiles à mesurer et à évaluer dans le contexte d'enquêtes effectuées avec des budgets très limités, avec des délais de publication qui se font de plus en plus serrés sous l'effet de demandes croissantes de la part de sociétés avides d'informations. Dans de telles conditions, on comprend que la priorité soit toujours donnée à la prévention plutôt qu'à la mesure et à la compensation, mais cela ne laisse guère de place à l'évaluation du succès des efforts de prévention, et réduit de ce fait les perspectives d'améliorations futures.

43. Certains utilisateurs qui connaissent mal les questions de statistiques peuvent donner une interprétation erronée aux informations sur les erreurs autres que les erreurs d'échantillonnage. C'est pourquoi la publication de ce type d'informations est parfois jugée non souhaitable pour certaines enquêtes, en raison de l'insuffisance des connaissances et d'une culture des statistiques, dont le développement peut s'avérer particulièrement difficile parmi des populations sachant insuffisamment lire et compter, comme c'est souvent le cas dans les pays en développement. Il est vrai également que, souvent, les organismes chargés des enquêtes n'ont pas de compétences suffisantes en statistiques, de sorte qu'il leur est difficile de cerner les problèmes et de prendre les mesures qui s'imposent pour y remédier, ainsi que de mesurer le succès de ces mesures. Quoi qu'il en soit, nous encourageons les organismes de statistiques à préparer et à publier de telles informations et à les rendre aussi claires et accessibles que possible aux adultes sachant lire.

44. Même si la situation est loin d'être satisfaisante, on note quelques signes encourageants. L'attention portée depuis peu à la question de la qualité des données par plusieurs grandes institutions de statistique, associations universitaires de statistique et d'enquête et même par des institutions publiques multilatérales est un élément positif. Les grandes initiatives auxquelles nous nous référons ici sont le Système général de diffusion de données (SGDD) et la Norme spéciale de diffusion de données (NSDD) du Fonds monétaire international (FMI), qui cherchent à promouvoir la normalisation de l'information sur la qualité des données statistiques par l'adhésion volontaire des pays à l'une ou l'autre de ces initiatives. D'après le FMI (2001), « Le SGDD est un processus structuré par lequel les pays membres du Fonds s'engagent d'eux-mêmes à améliorer à long terme la qualité des données produites et diffusées par leurs systèmes de statistiques afin de répondre aux besoins de l'analyse macroéconomique ». Toujours d'après le FMI : « Le SGDD favorise de saines pratiques statistiques pour l'établissement et la diffusion de statistiques économiques, financières et sociodémographiques. Il permet

d'identifier des ensembles de données particulièrement utiles pour l'analyse économique et le suivi des phénomènes sociaux et démographiques, et de fixer des objectifs et formuler des recommandations pour leur conception, leur production et leur diffusion. Il accorde une attention particulière aux besoins des utilisateurs, qu'il s'efforce de satisfaire en suivant des directives concernant la qualité et l'intégrité des données et l'accès du public à ces données » (ibid).

45. La principale contribution de ces initiatives est d'offrir aux pays : *a*) un cadre pour la qualité des données (voir <http://dsbb.imf.org/dqrsindex.htm>) qui facilite l'identification des principales sources de problèmes et d'objectifs pour l'amélioration de la qualité des données; *b*) l'incitation économique à envisager d'améliorer la qualité des données d'un ensemble d'enquêtes et de travaux statistiques (en leur offrant ou en leur renouvelant l'accès aux marchés financiers internationaux); *c*) une incitation commune ou partagée à favoriser le débat sur la qualité des données sans crainte d'erreurs d'interprétation; et *d*) s'il y a lieu, un appui technique aux programmes d'évaluation et d'amélioration. Il ne s'agit pas d'une initiative mondiale, car tous les pays ne sont pas membres du FMI. Toutefois, 131 pays ont été contactés à ce sujet, et à une date récente, 46 d'entre eux avaient décidé d'adhérer au SGDD et 50 autres pays avaient atteint le statut plus élevé d'abonnés à la NSDD, après avoir satisfait à un ensemble de contrôles et de critères plus rigoureux d'évaluation de la qualité de leurs données statistiques.

46. L'examen détaillé des normes de qualité des données encouragées par le FMI et d'autres organisations n'entre pas dans le cadre du présent chapitre, mais le lecteur est encouragé à approfondir cette question en se reportant aux références données ici. Les pays en développement devraient s'associer au débat actuel sur les normes, décider d'adhérer ou non à l'une ou l'autre des initiatives susmentionnées et, éventuellement, contribuer à la définition ou à la révision de ces normes. Mais surtout, les organismes statistiques des pays en développement peuvent utiliser ces normes comme point de départ (s'ils n'ont rien d'approchant chez eux) pour promouvoir la recherche d'une plus grande qualité chez leurs membres et leur personnel et parmi les utilisateurs.

47. L'autre initiative que nous mentionnerons ici, en particulier parce qu'elle intéresse le Brésil et d'autres pays d'Amérique latine, est le Projet de coopération statistique entre l'Union européenne (UE) et le Marché commun du Sud (MERCOSUR¹). D'après l'objectif de ce projet : « L'Union européenne et les pays du MERCOSUR ont signé un accord de « Coopération statistique avec les pays du MERCOSUR », dont le principal but est le rapprochement² dans les méthodes statistiques afin de permettre l'utilisation de diverses données statistiques selon des termes convenus mutuellement, en particulier dans le domaine des échanges de biens et de services et, d'une façon générale, dans tout domaine se prêtant à la mesure statistique ». Ce projet « devrait permettre à la fois la normalisation des méthodes statistiques entre les pays du MERCOSUR et entre ces pays et l'Union européenne ». (Pour plus de détails, consulter le site Internet : <http://www.ibge.gov.br/mercosur/english/index.html>.) Il a déjà permis de promouvoir un certain nombre de cours et stages de formation et, ce faisant, contribue à améliorer la pratique des enquêtes et une plus grande sensibilité aux erreurs et à leur mesure.

48. Les initiatives comme celle-ci sont essentielles pour améliorer la position des organismes de statistiques des pays en développement : leurs statistiques peuvent être de bonne qualité, mais souvent, ils ne savent pas combien elles sont bonnes. La coopération internationale des pays développés avec les pays en développement et entre ces derniers est essentielle au progrès de la mesure et de la notification des erreurs

¹ MERCOSUR est le marché commun du Sud, formé d'un groupe de pays comprenant l'Argentine, le Brésil, le Paraguay et l'Uruguay, qui partagent un accord de libre-échange.

² Ce terme est utilisé ici dans le sens d'harmonisation.

autres que les erreurs d'échantillonnage ainsi qu'à d'autres aspects de la qualité des données statistiques.

D. Autres lectures recommandées

49. Il est recommandé de lire les écrits sur les réunions suivantes :

- Conférence internationale sur la mesure des erreurs dans les enquêtes, tenue à Tucson, Arizona en 1990 (voir Biemer *et al.*, 1991).
- Conférence internationale sur la mesure dans les enquêtes et sur la qualité du processus d'enquête, tenue à Bristol, Royaume-Uni, en 1995 (voir Lyberg *et al.*, 1997).
- Conférence internationale sur la non-réponse dans les enquêtes, tenue à Portland, Oregon, en 1999 (voir Groves *et al.*, 2001).
- Conférence internationale sur la qualité des statistiques officielles, tenue à Stockholm, Suède, en 2001 (visiter <http://www.q2001.scb.se/>).
- Symposium 2001 de Statistique Canada, tenu à Ottawa, Canada, sur la qualité des données d'une agence de statistique d'un point de vue méthodologique (visiter <http://www.statcan.ca/english/conferences/symposium2001/session21/s21c.pdf>).
- Cinquante-troisième session de l'Institut international de statistique (IIS), tenue à Séoul, République de Corée, en 2001, où a été présentée une étude sur « les programmes de qualité dans les agences de statistiques », traitant des approches de la qualité des données des offices nationaux et internationaux de statistique (visiter <http://www.nso.go.kr/isi2001>).
- Séminaire 2000 sur la qualité des statistiques, parrainé par le FMI, tenu à l'île de Jeju, République de Corée, en 2000 (visiter <http://www.nso.go.kr/sqs2000/sqs12.htm>).
- Conférence internationale sur l'amélioration des enquêtes, tenue à Copenhague, Danemark, en 2002 (visiter <http://www.icis.dk/>).

RÉFÉRENCES

- Bianchini, Z. M., et S. Albieri (1998). A review of major household sample survey designs used in Brazil. Dans *Proceedings of the International Conference on Statistics for Economic and Social Development*. Aguascalientes, Mexique, 1998, Instituto Nacional de Estadística, Geografía e Informática (INEGI).
- Biemer, P. P., et R. S. Fecso (1995). Evaluating and controlling measurement error in business surveys, Cox *et al.*, eds. Dans *Business Survey Methods*, New York, John Wiley and Sons.
- Biemer, P. P., *et al.* (1991). *Measurement Errors in Surveys*. New York, John Wiley and Sons.
- Correa, S. T., P. L. do Nascimento Silva and M. P. S. Freitas (2002). Estimación de variância para o estimador da diferença entre duas taxas na pesquisa mensal de emprego. Dans *15^o Simpósio Nacional de Probabilidade e Estatística*. Aguas de Lindóia, Brésil, São Paulo, Brésil, Associação Brasileira de Estatística.
- Criado, I. V., et M. S. B. Cabria (1990). *Procedimiento de depuración de datos estadísticos*, cuaderno 20. Vitoria-Gasteiz, Espagne : EUSTAT Instituto Vasco de Estadística.

- Fonds monétaire international (2001). Guide to the General Data Dissemination System (GDDS). Washington, D.C., IMF Statistics Department. Accessible à <http://dsbb/imf/org/applications/web/gdds/gddsguidelangs>.
- Garcia Rubio, E., et I. V. Criado (1990). DIA System: software for the automatic imputation of qualitative data. Dans *Proceedings of the United States Census Bureau Sixth Annual Research Conference* (Arlington, Virginie). Washington, D.C., United States Bureau of the Census.
- Groves, R. M., et al. (2001). *Survey Non-response*. New York, John Wiley and Sons.
- Instituto Brasileiro de Geografia e Estatística (2002a).
<http://www.ibge.gov.br/home/estatistica/populacao/trabalhoerendimento/pnad99/metodologia99.shtm>.
- _____ (2002b). <http://www.ibge.net/home/estatistica/indicadores/trabalhoerendimento/pme/default.shtm>.
- Kalton, G. (1983). *Compensating for Missing Survey Data*. Research Report Series. Ann Arbor, Michigan, Institute for Social Research, Université du Michigan.
- _____ (1986). Handling wave non-response in panel surveys. *Journal of Official Statistics*, vol. 2, n° 3, pp. 303-314.
- _____ (2002). Models in the practice of survey sampling (revisited). *Journal of Official Statistics*, vol.18, n° 2, pp. 129-154.
- Kordos, J. (2002). Personal communications.
- Lyberg, L., et al., eds. (1997). *Survey Measurement and Process Quality*. New York, John Wiley and Sons.
- Marks, E. S. (1973). The role of dual system estimation in census evaluation. Internal report. Washington, D.C., United States Bureau of the Census.
- Nations Unies (1982). *National Household Survey Capability Programme: Non-sampling errors in household surveys: sources, assessment and control: Preliminary Version*. DP/UN/INT-81-041/2. New York, Département de la coopération technique au développement et Bureau de statistique.
- Oliveira, L. C., et al. (2003). *Censo Demográfico 2000: Resultados da Pesquisa de Avaliação da Cobertura da Coleta*. Textos para Discussão, n° 9. Rio de Janeiro, IBGE, Directoria de Pesquisas.
- Pfeffermann, D., P. L. Nascimento de Silva et M. P. S. Freitas (2000). Implications of the Brazilian Labour Force rotation scheme on the quality of published estimates. Internal report. Rio de Janeiro, IBGE, Departamento de Metodologia.
- Platek, R., et C. E. Särndal (2001). Can a statistician deliver? *Journal of Official Statistics*, vol. 17, n° 1, pp. 1-20.
- Poirier, P., M. Bankier et M. Lachance (2001). Efficient methodology within the Canadian Census Edit and Imputation System (CANCEIS). Paper presented at the Joint Statistical Meetings, American Statistical Association.
- Särndal, C. E., B. Swensson et J. Wretman (1992). *Model Assisted Survey Sampling*. New York, Springer-Verlag.

Section D

Coûts des enquêtes

Introduction

JAMES LEPKOWSKI

Université du Michigan
Ann Arbor, Michigan
États-Unis d'Amérique

1. Dans les précédentes sections, nous avons examiné les erreurs d'échantillonnage et les erreurs autres que les erreurs d'échantillonnage afin de mieux saisir ce que l'on entend par la qualité des estimations des enquêtes. Presque toutes ces erreurs se prêtent à des méthodes pour réduire leur ampleur. Cependant, l'application de ces méthodes comporte souvent un coût additionnel. Comme les enquêtes disposent d'un budget fixe pour couvrir leurs dépenses, pour consacrer des ressources supplémentaires à combattre une source d'erreurs, il faut transférer des ressources d'un domaine à un autre. La conception des enquêtes suppose de constants arbitrages entre les considérations de coût et d'erreur.

2. Par exemple, supposons que, dans une enquête particulière sur les ménages, un sous-groupe de population parle une langue dans laquelle le questionnaire n'a pas été traduit. Les concepteurs de l'enquête peuvent décider d'abord de réduire la taille de l'échantillon pour diminuer les coûts de l'enquête, puis utiliser l'économie ainsi réalisée pour traduire le questionnaire dans une langue supplémentaire ou pour engager des enquêteurs qui parlent cette langue, et étendre ainsi l'enquête à ces ménages.

3. Étant donné que la conception d'une enquête comporte souvent une série de compromis de ce type, pour pouvoir décider en connaissance de cause, il faut disposer de bonnes informations sur la nature et l'ampleur des erreurs émanant de différentes sources (telles que la variance de l'échantillonnage et la distorsion engendrée par la non-couverture, dans l'exemple précédent) et sur les coûts que comportent différentes procédures d'enquête. Dans la précédente section, nous avons examiné les sources d'erreurs et l'ampleur des erreurs. Dans la présente section, nous examinerons la nature des coûts des enquêtes.

4. Dans une enquête, les considérations de coût se posent à trois stades. Le premier est celui de la planification de l'enquête, où les coûts sont estimés d'avance. Les estimations de coût de la planification ou « phase d'établissement du budget » sont difficiles à établir, à moins que l'on puisse s'inspirer d'une expérience précédente. Les enquêtes continues peuvent fournir des données pertinentes sur les coûts pour la planification des cycles suivants, quoique ces données se heurtent souvent aux considérations de coût au stade suivant – le suivi des coûts des enquêtes.

5. Les organismes d'enquête et certaines autres institutions qui, à l'occasion, effectuent des enquêtes disposent rarement de systèmes perfectionnés susceptibles de leur permettre de suivre les coûts de manière à pouvoir utiliser les données ainsi obtenues pour la planification. Ces coûts sont réunis dans un système comptable, mais ce système ne classe pas les coûts selon les types de catégories dont les concepteurs ont besoin pour leur planification. Lorsqu'un tel suivi des coûts est tenté, il risque d'alour-

dir le coût de l'enquête proprement dite s'il nécessite l'adjonction de nouveaux systèmes aux opérations.

6. Si l'on suit les coûts dans une opération en cours, on peut envisager d'apporter plus systématiquement des modifications de conception pendant la collecte des données. On peut utiliser les données sur les coûts pour prédire l'ampleur de l'économie réalisable dans le cadre d'une opération et l'impact de la réaffectation éventuelle des ressources à un autre domaine.

7. On décide de la réaffectation des ressources lors de la planification de l'enquête en étudiant les arbitrages entre le niveau de coût et les erreurs pouvant provenir de sources multiples. Ces arbitrages, qui peuvent se faire et se font notamment au stade de la conception de l'échantillon, permettent de trouver une solution optimale au problème de l'affectation des ressources.

8. En effet, comme on l'a vu au chapitre II, pour les enquêtes qui se font sur des grappes prélevées sur un échantillon probabiliste pris lui-même sur une population très éparse, il importe de limiter le nombre de grappes pour réduire les coûts de collecte des données. Toutefois, quand on limite le nombre de grappes, il faut accroître le nombre d'observations faites sur chaque grappe, afin de maintenir la taille globale de l'échantillon. Or, cet accroissement de la taille du sous-échantillon de chaque grappe accentue la variabilité des estimations. Autrement dit, à mesure que les coûts diminuent sous l'effet de la diminution du nombre de grappes, la variance de l'échantillonnage augmente. Ce qu'il faut, ce sont des indications sur le nombre de grappes à choisir de manière à limiter les coûts, compte tenu du niveau de précision requis et de la nécessité de maintenir la variance aussi faible que possible pour un niveau de coût donné. Pour la conception de l'échantillon, on dispose d'une solution mathématique à ce problème.

9. L'arbitrage coût-erreur apparaît également à d'autres niveaux de la conception de l'échantillon. Par exemple, on peut réduire le taux de non-réponse à une enquête sur les ménages en rendant plusieurs visites aux ménages qui ne répondent pas lors de la première visite. On peut demander à l'enquêteur d'effectuer jusqu'à quatre ou cinq visites à un même ménage pour en obtenir une réponse. Cependant, le coût de ces visites répétées oblige à limiter la taille de l'échantillon. Autrement dit, le coût des efforts de réduction de la distorsion liée à la non-réponse accentue la variance de l'échantillonnage. Là encore, ces efforts d'allègement des coûts dans un domaine oblige à un redéploiement des ressources et engendre la possibilité de nouvelles erreurs dans un autre domaine de la conception de l'enquête.

10. Dans les chapitres de cette section, nous examinerons un certain nombre de questions concernant la planification, le suivi et le redéploiement des coûts de la conception d'une enquête. Nous utiliserons les données d'enquêtes sur les ménages effectuées dans des pays en développement et des pays en transition pour illustrer les types de coûts que comporte la collecte des données et, dans une certaine mesure, leur ampleur. Comme les enquêtes varient considérablement d'un pays à un autre et plus encore d'un continent à un autre, les informations précises sur les coûts peuvent se révéler peu utiles pour la planification d'une enquête dans un pays donné. On espère toutefois que les sources et les niveaux de ces coûts présentés dans les chapitres suivants aideront les concepteurs opérant dans des cadres très différents à mieux comprendre ce qu'impliquent ces coûts et les arbitrages entre coûts et erreurs pour leurs propres enquêtes.

Chapitre XII

Analyse des questions de coût concernant les enquêtes effectuées dans les pays en développement et les pays en transition

IBRAHIM S. YANSANEH*

Commission de la fonction publique internationale
Nations Unies, New York

* Ancien chef, Service de la méthodologie et de l'analyse, Division de statistique, Nations Unies.

RÉSUMÉ

Le présent chapitre traite, en termes généraux, des principales questions concernant le coût de conception et d'exécution d'enquêtes sur les ménages dans les pays en développement et les pays en transition. Le coût total de l'enquête est décomposé en plusieurs éléments liés aux divers aspects de sa conception et de son exécution. Les facteurs de coût sont examinés successivement pour les pays disposant d'une vaste infrastructure pour les enquêtes et pour ceux pour lesquels cette infrastructure est peu développée ou inexistante. Nous examinerons également la comparabilité des coûts selon les pays.

Termes clés : infrastructure utilisée pour les enquêtes, coût supplémentaire par entrevue, efficacité, comparabilité des coûts, facteurs de coût.

A. Introduction

1. Critères d'efficacité pour la conception d'échantillons

1. D'une façon générale, pour être efficace, un échantillon doit répondre à deux critères : il doit permettre d'établir des estimations raisonnablement précises dans les limites d'un budget fixe ou limiter le coût d'obtention d'un niveau de précision donné. Le présent chapitre se concentre sur le premier de ces critères, qui concerne la conception de l'échantillon le plus efficace dont l'application soit compatible avec le budget disponible et de nature à permettre une utilisation raisonnablement efficace des ressources. Dans les pays en développement et les pays en transition, le coût des enquêtes est l'un des principaux obstacles à la formulation de décisions critiques de conception et d'exécution. Dans ces pays comme dans les pays développés, la conception d'une enquête implique les arbitrages habituels entre la précision des estimations et le coût d'exécution. La précision se mesure généralement aux variances des estimateurs des grandeurs considérées comme présentant le plus d'intérêt pour une population donnée. Parmi les autres mesures de la précision figurent le carré de l'erreur moyenne ou l'erreur totale, qui comprend également l'élément de distorsion de l'erreur.

2. Le développement mathématique formel des arbitrages entre la précision et le coût suppose généralement l'optimisation de la variance ou des fonctions de coût

sujettes à des contraintes relativement simples. Toutefois, en raison de la limitation des informations sur le coût et la variance, cette approche de l'optimisation ne doit souvent être perçue que comme offrant une approximation grossière de la conception optimale ou de la précision et du coût effectifs d'exécution. Ces questions ont fait l'objet d'un examen approfondi pour les enquêtes effectuées dans les pays en développement. [Voir, par exemple, Andersen, Kasper et Frankel (1979), Cochran (1977), Groves (1989), Kish (1965; 1976) et Linacre et Trewin (1993), et les références que citent ces auteurs. Par ailleurs, pour un examen plus général du coût et de la précision comme deux des principaux critères d'évaluation des systèmes nationaux de statistiques, voir De Vries (1999, p. 70) et les références citées dans son étude. Pour des analyses empiriques des coûts de certaines enquêtes dans les pays en développement et les pays en transition, voir chapitres XIII et XIV et l'introduction à la section D (Coûts des enquêtes).]

3. L'une des principales limitations dont souffre la conception des enquêtes dans les pays en développement et les pays en transition tient au manque ou à l'insuffisance d'informations sur les coûts liés aux divers aspects de l'exécution des enquêtes. Malgré ces limitations, on trouve souvent certains éléments structurels communs dans les coûts des enquêtes, qui peuvent être utiles pour la conception d'une nouvelle enquête. Dans certains cas, cette structure commune se limite à des indications qualitatives de l'ampleur relative de plusieurs éléments de coût ou à leurs sources. Dans d'autres cas, on dispose des coûts effectifs qui peuvent être perçus comme relativement homogènes parmi un ensemble de pays, surtout de pays ayant plus ou moins la même distribution démographique et les mêmes niveaux d'infrastructure utilisée pour les enquêtes.

4. Ce chapitre présente une analyse des questions de coût des enquêtes dans les pays en développement et les pays en transition et étudie dans quelle mesure les coûts des enquêtes ou certains éléments connexes observés dans un pays peuvent servir à améliorer la conception d'enquêtes similaires dans un autre pays. Autrement dit, il s'efforce de traiter la question de la transférabilité des coûts des enquêtes d'un pays à un autre. Une telle analyse est utile à deux titres : premièrement, elle peut apporter une solution partielle au problème de l'insuffisance d'informations sur le coût des enquêtes dans les pays en développement et les pays en transition; et deuxièmement, dans la mesure où on note des similarités entre pays en ce qui concerne la conception d'échantillons, l'infrastructure utilisée pour les enquêtes et la distribution de la population, on peut s'attendre à trouver des similarités, ne serait-ce que dans certaines éléments du coût des enquêtes dans ces pays. On peut utiliser des informations sur les coûts d'une enquête effectuée dans un pays pour concevoir une enquête dans un autre pays ou pour améliorer l'efficacité de conception des enquêtes dans le premier pays. Pour ce faire, le concepteur doit tenir compte de la grande variabilité des structures de coût des enquêtes selon les pays. Les éléments de coût variables sont généralement spécifiques au pays, tandis que les coûts fixes ont des chances d'être comparables d'un pays à un autre.

2. Éléments de la structure des coûts des enquêtes dans les pays en développement et les pays en transition

5. Dans ce chapitre, nous nous concentrons sur le premier critère d'efficacité de conception des enquêtes, à savoir une conception qui produit des estimations raisonnablement précises dans les limites d'un budget donné. De nombreuses enquêtes sont effectuées dans les pays en développement et les pays en transition à la demande d'institutions financières internationales qui ont besoin de données pour décider de

participer à certains projets de développement ou pour appuyer les décideurs et les pouvoirs publics des pays bénéficiaires. Les Enquêtes démographiques et sanitaires (DHS) effectuées par ORC Macro pour l'United States Agency for International Development; l'Enquête sur la mesure des niveaux de vie (LSMS) effectuée par la Banque mondiale; et les Enquêtes en grappes à indicateurs multiples (MICS) effectuées par le Fonds des Nations Unies pour l'enfance (UNICEF) sont trois des principaux exemples d'enquêtes effectuées dans les pays en développement et les pays en transition. Par ailleurs, de nombreuses autres enquêtes sont effectuées régulièrement par les offices nationaux de statistiques et d'autres organismes des systèmes nationaux de statistiques. Il y a également un grand nombre d'enquêtes de plus petite envergure effectuées à la demande de donateurs par de petites organisations locales (par exemple, des organisations non gouvernementales). Bien entendu, la question du coût est tout aussi critique dans le travail de conception de ces enquêtes.

6. S'agissant des questions de coût, il importe de tenir compte du fait que les enquêtes effectuées dans les pays en développement ont beaucoup de traits communs. Par exemple, la plupart d'entre elles sont fondées sur un échantillon probabiliste stratifié à plusieurs degrés. Les unités primaires d'échantillonnage (UPE) sont souvent établies à partir de zones d'énumération identifiées et utilisées lors d'un précédent recensement national de la population. Les unités secondaires d'échantillonnage (USE) sont généralement des unités d'habitation ou des ménages et les unités ultimes d'échantillonnage (UUE) sont généralement des ménages ou des personnes. Les strates et les domaines analytiques sont généralement le produit de l'intersection de régions administratives et de sous-domaines urbains/ruraux de ces régions. En raison de ces similarités, et conformément aux études mentionnées ci-dessus au paragraphe 2, il peut être utile d'étudier dans quelle mesure on peut identifier des structures de coût communes dans des groupes d'enquêtes effectuées dans des pays en développement. [Pour plus d'informations générales sur la conception et l'exécution d'enquêtes dans les pays en développement et les pays en transition, voir section A de la première partie (Conception et mise en œuvre des enquêtes) et les études de cas de la partie deux de cette publication. Pour un examen plus approfondi des éléments de coût d'une enquête donnée dans un pays en développement, voir chapitre XIII.] Des comparaisons empiriques des éléments de coût des enquêtes effectuées dans certains pays en développement et certains pays en transition sont présentées aux chapitre XIV.

7. Dans ce chapitre, nous nous limiterons à l'examen des grandes enquêtes nationales sur les ménages effectuées par les offices nationaux de statistiques et autres organismes publics du système national de statistiques. Parmi celles-ci figurent notamment les enquêtes sur le budget, les revenus et les dépenses des ménages et les enquêtes démographiques et sanitaires. Bien que les enquêtes sur les marchés et autres enquêtes plus modestes sur les ménages effectuées de façon ad hoc par divers organismes constituent une source d'informations utile qui alimente les décisions de politique nationale et les plans de développement, elles ne sont pas prises en compte dans cette analyse. Néanmoins, les grandes questions soulevées ici s'appliquent également à ces enquêtes. La plupart des exemples sont tirés des Enquêtes démographiques et sanitaires (DHS) et des Enquêtes sur la mesure des niveaux de vie (LSMS), mais les questions clés s'appliquent de façon générale à toutes les enquêtes sur les ménages.

3. Aperçu général de ce chapitre

8. Ce chapitre s'articule de la façon suivante : la section B traite de la décomposition classique du coût global de l'enquête en diverses composantes. Les trois sections suivantes donnent une description qualitative de certains facteurs qui influent sur les

coûts globaux des enquêtes effectuées dans les pays en développement et les pays en transition. La section C analyse les facteurs de coût qui peuvent être importants dans les cas où une infrastructure considérable est déjà en place pour les enquêtes. La section D porte sur les cas où l'infrastructure est limitée ou inexistante. La section E traite des modifications de la structure des coûts résultant des modifications apportées aux objectifs des enquêtes. La section F contient une mise en garde concernant l'interprétation des coûts annoncés des enquêtes. La section G présente certaines conclusions et un résumé de certains points saillants qui n'ont pas été pleinement développés dans la discussion. Un exemple de cadre utilisé pour l'établissement du budget des enquêtes en grappes à indicateurs multiples (MICS) effectuées par l'UNICEF dans les pays en développement et les pays en transition est décrit en annexe par Ajayi (2002).

B. Composantes du coût des enquêtes

9. L'analyse mathématique des coûts des enquêtes postule généralement un coût global, C , comme fonction linéaire des nombres d'unités primaires d'échantillonnage choisies et de certains éléments. L'équation ci-dessous offre un exemple de cette fonction :

$$C = c_0 + \sum_{b=1}^L n_b c_b + \sum_{b=1}^L \sum_{i=1}^{n_b} n_{bi} c_{bi} \quad (1)$$

où c_0 représente les coûts fixes de lancement de l'enquête, c_h est le coût additionnel de collecte d'informations auprès d'une unité primaire d'échantillonnage (UPE) supplémentaire dans une strate h ; n_h est le nombre d'UPE sélectionnées, c_{hi} est le coût additionnel d'entrevue d'un ménage supplémentaire au sein de l'UPE i de la strate h ; et n_{hi} est le nombre de ménages sélectionnés au sein de l'UPE i . [Voir, par exemple, Cochran (1977, sect. 5.5 et 11.13-11.14) et Groves (1989, chap. 2).] En général, les coefficients de coût c_0 , c_h et c_{hi} dépendent d'un grand nombre de facteurs qui peuvent varier selon les pays et selon les enquêtes effectuées dans un même pays. Ces facteurs sont examinés de façon détaillée dans les sections suivantes.

10. L'équation (1) n'est que l'une des nombreuses expressions possibles du coût. Ainsi, Cochran (1977, p. 313) examine l'inclusion d'une composante de coût distincte liée à l'utilisation d'une liste d'unités secondaires d'échantillonnage (comme phase intermédiaire avant le sous-échantillonnage des ménages pour les entrevues) parmi les unités primaires d'échantillonnage sélectionnées, où cette composante est fonction du nombre d'unités secondaires de chaque unité primaire. En outre, pour un échantillon à trois phases, où les personnes sont choisies au hasard au sein des ménages pour l'entrevue, il faut ajouter un terme à l'équation (1) ci-dessus pour le coût additionnel lié à l'entrevue avec une personne supplémentaire au sein d'un ménage sélectionné.

11. Par ailleurs, une fonction de coût par degrés est souvent préférable à une fonction linéaire. En effet, si l'on peut effectuer 10 entrevues en une journée, si l'on procède à une onzième entrevue, il faut rajouter un jour de travail et, par conséquent, un coût supplémentaire élevé, alors qu'en ajoutant une douzième entrevue, on n'alourdit que légèrement le coût global. Il importe également de noter que les décisions telles que celles portant sur le nombre d'UPE sont parfois fonction de considérations pratiques autres que les considérations de coût et de précision. Par exemple, on peut imaginer que l'on veuille consacrer une semaine entière aux entrevues dans une UPE. Dans ce cas, on ne peut envisager une charge de travail inférieure à celle d'une semaine, alors qu'une double charge de travail correspondant à deux semaines de travail serait envisageable. Autrement dit, en pareil cas, le nombre d'UPE ne serait pas directement

fonction des considérations de coût et des effets liés à la conception, mais des difficultés pratiques d'exécution.

12. Dans la section suivante, nous examinerons les coûts d'exécution des enquêtes en fonction du niveau de l'infrastructure dont dispose le pays à cette fin, l'idée étant qu'il y a une énorme disparité entre les coûts globaux des enquêtes selon que les pays sont pourvus d'une riche infrastructure ou, au mieux, d'une infrastructure insuffisante. Cela dit, dans les pays en développement comme dans les pays en transition, il importe d'évaluer le degré d'infrastructure dès le stade de la planification de l'enquête et de ne pas compter sur les données dont on dispose sur les antécédents. Il n'est pas rare qu'un pays doté d'une excellente infrastructure voie cette infrastructure se détériorer peu à peu, au point qu'il doive redescendre du premier groupe de pays (examiné à la section C) au second (examiné à la section D).

C. Coûts des enquêtes effectuées dans les pays disposant d'une vaste infrastructure

1. Facteurs liés aux activités préparatoires

13. Une large part du coût d'une enquête ponctuelle est consacrée au financement des activités préparatoires [voir, par exemple, Grosh et Muñoz (1996, p. 199)], et les fonds destinés à ces activités sont décaissés au début du processus d'enquête. Les coûts sont relativement fixes pour certaines activités préparatoires telles que la coordination de la planification de l'enquête par plusieurs organismes publics, l'élaboration du cadre d'échantillonnage, la conception de l'échantillon, la conception du questionnaire, l'impression de ce questionnaire et d'autres documents en rapport avec l'enquête et la publicité à l'intention des répondants potentiels. En revanche, certains coûts, tels que les coûts de recrutement et de formation du personnel de terrain (par exemple, des personnes chargées d'établir les listes, des enquêteurs, des superviseurs et des traducteurs) sont fonction de la taille de l'échantillon (que ce soit au niveau des unités primaires ou des ménages).

14. Les coûts des activités préparatoires dépendent de certains facteurs locaux, tels que l'ampleur du personnel chargé de l'enquête et son niveau de rémunération, le type et la quantité de matériel, les prix de certains articles tels que les fournitures de bureau et autres et les moyens de transport et de communication. De plus, les coûts sont très différents selon que l'enquête est une étude transversale effectuée pour la première fois, dans ce cas, les coûts unitaires sont relativement plus élevés, ou qu'il s'agit d'une enquête permanente, pour laquelle les coûts unitaires sont plus faibles.

2. Facteurs liés à la collecte et au traitement des données

15. Les coûts de collecte et de traitement des données comprennent aussi des éléments fixes et des éléments variables, mais les coûts de collecte des données sont pour la plupart variables et sont fonction du nombre d'unités primaires d'échantillonnage et de ménages sélectionnés. Ces coûts portent notamment sur l'établissement des listes de ménages au sein des unités primaires d'échantillonnage ou des listes de personnes au sein des ménages sélectionnés et sur les entrevues et la supervision sur le terrain. Les coûts de collecte des données comprennent aussi les coûts de déplacement, tant d'une UPE à une autre qu'à l'intérieur des UPE. Ces coûts sont fonction de l'organisation des entrevues, de la longueur du questionnaire, de la nécessité éventuelle d'utiliser des interprètes et du nombre d'unités à interroger.

16. On peut réduire les coûts des déplacements en créant des équipes nationales composées de superviseurs et d'enquêteurs que l'on déplace d'une région à une autre, plutôt que de former des équipes régionales. Cette option offre en outre l'avantage d'améliorer la qualité des données. Cette formule peut aussi être utile lorsque la collecte des données se fait de façon continue, ou lorsque l'enquête nécessite un matériel coûteux. La solution des équipes multiples est utilisée pour de nombreuses enquêtes effectuées dans les pays en développement et les pays en transition, notamment pour les Enquêtes sur la mesure des niveaux de vie (LSMS) [Grosh et Muñoz, 1996, chap. 5]. Dans ceux de ces pays où les langues diffèrent d'une région à l'autre, il peut être préférable que les équipes soient formées de personnes parlant la langue de la région.

17. Les coûts de coordination des activités de terrain et d'établissement de la documentation de l'enquête représentent une part importante des coûts de collecte et de traitement des données. Avec un système centralisé de collecte et de traitement des données, les coûts liés au regroupement des questionnaires remplis et à leur envoi au siège peuvent être substantiels. Par ailleurs, le budget doit tenir compte des coûts potentiellement élevés liés au suivi des activités et des résultats de l'enquête, par exemple, aux procédures d'établissement des listes et de sous-échantillonnage sur le terrain, et aux taux de réponse pour certains domaines clés mesurés par rapport à des niveaux prédéterminés, etc. Lorsque le suivi de ces activités est assuré efficacement, les personnes chargées d'effectuer l'enquête peuvent, s'il y a lieu, prendre des mesures rectificatives pendant la collecte des données plutôt que d'avoir à constater les insuffisances après coup, auquel cas le coût des mesures rectificatives pourrait être prohibitif.

18. Les travaux d'entrée, d'édition et d'imputation des données, qui font partie du traitement des données, peuvent comporter un ensemble de coûts fixes et de coûts variables, selon le degré d'automatisation de ce processus. Les autres principaux coûts de traitement des données sont généralement fixes et comprennent le coût d'achat des ordinateurs et du logiciel, d'établissement des pondérations et d'estimation des variances ainsi que des autres travaux d'analyse des données. Par exemple, les pondérations seraient calculées indépendamment du nombre d'UPE ou de ménages sélectionnés; et une fois qu'une procédure de pondération aurait été établie et programmée, le coût additionnel de calcul d'une pondération pour un ménage supplémentaire serait négligeable.

19. Le coût de traitement des données dépend du nombre de niveaux d'analyse prévus au budget. Pour certaines enquêtes, seule une analyse préliminaire est effectuée sur les données recueillies présentées en tableau. Pour d'autres enquêtes comme les DHS et les LSMS, des analyses statistiques plus détaillées sont effectuées sur la base de recommandations à l'intention des gouvernements bénéficiaires et des institutions d'aide. Ainsi, les DHS et les LSMS comportent divers types d'analyse détaillée de leurs microdonnées et publient leurs conclusions dans une série de rapports (dans le cas des DHS) et de documents de travail (dans celui des LSMS). Certaines des références citées ci-dessous en offrent des exemples. Les coûts de production de rapports et de diffusion des résultats, de même que ceux de divers services nécessaires aux analystes, tels que la préparation de métadonnées et l'organisation d'ateliers de formation, sont considérables.

D. Coûts des enquêtes dans les pays où l'infrastructure nécessaire est insuffisante ou inexistante

20. Dans les pays où l'infrastructure nécessaire aux enquêtes est relativement rudimentaire, il y a peu de chances que l'organisme parrainant l'enquête consacre au

renforcement des capacités un volume substantiel de ressources que n'aurait pas à investir un pays doté d'une solide infrastructure (Grosh et Muñoz, 1996, chap. 8). Le manque d'infrastructure peut alourdir considérablement les coûts des travaux préparatoires, des opérations sur le terrain et du traitement des données.

21. Le renforcement des capacités implique généralement une vaste formation initiale du personnel. Dans un pays où l'infrastructure nécessaire aux enquêtes est limitée ou absente, le recours à des experts extérieurs pour la conception de l'enquête est généralement coûteux. De plus, le temps du personnel de terrain tend à être utilisé plus efficacement à mesure que l'organisme enquêteur acquiert de l'expérience. Qui plus est, dans les pays possédant une vaste expérience des enquêtes, le besoin de déplacement est nettement inférieur car les offices de statistiques ont des chances de disposer d'équipes régionales de collecte de données expérimentées ou de fournir des moyens de transport au personnel de terrain. Ces avantages se traduisent par des économies sur les coûts de transport, de formation et autres coûts de personnel. Les pays n'ayant pas l'expérience d'enquêtes précédentes prévoient généralement des véhicules dans le budget de l'enquête, et cet élément peut devenir l'un des principaux postes de ce budget (Grosh et Muñoz, 1996, chap. 8). Le matériel informatique et les cartes servant à l'identification des ménages sont d'autres exemples de postes budgétaires sur lesquels l'existence d'une certaine infrastructure ou l'expérience d'enquêtes précédentes a un profond impact.

E. Facteurs liés aux modifications des objectifs de l'enquête

22. Comme on l'a indiqué précédemment, de nombreux facteurs influant sur les coûts sont liés à certains aspects de la conception de l'enquête, notamment la taille de l'échantillon, la longueur du questionnaire, le nombre de modules et certaines méthodes employées dans la sélection de l'échantillon et l'établissement des listes, l'expérimentation pilote et la conception et la traduction du questionnaire. Pour une conception donnée, certains des coûts sont plus ou moins constants d'un pays à un autre.

23. Cependant, dans les pays en développement et les pays en transition, la conception des enquêtes doit souvent être modifiée pour tenir compte de certaines spécifications spéciales des gouvernements bénéficiaires ou autres parties prenantes. Par exemple, le gouvernement peut décider d'élargir les objectifs de l'enquête et d'y inclure d'autres priorités nationales. Cela peut se traduire par : *a*) l'addition de nouveaux modules au questionnaire; ou *b*) un accroissement du nombre de domaines si l'on veut que les estimations des variables clés aient la même précision pour certains groupes sous-nationaux que pour l'ensemble du pays.

24. Ces modifications peuvent influencer de plusieurs façons sur les arbitrages entre le coût et la qualité des données. Premièrement, elles peuvent se solder par de forts accroissements de la durée totale nécessaire à l'enquêteur pour la collecte des données en raison de l'inclusion de modules additionnels au questionnaire [par. 23, *a*] ou en raison de l'accroissement exponentiel du nombre d'entrevues du fait de l'accroissement du nombre de domaines (par. 23, *b*). Deuxièmement, si l'organisme chargé de l'enquête dispose d'un nombre relativement fixe d'enquêteurs et de superviseurs qualifiés, les modifications peuvent obliger à former des enquêteurs supplémentaires et à prévoir davantage de temps de supervision par minute de temps d'entrevue, ou bien le nombre d'agents de terrain qualifiés peut être maintenu constant moyennant une prolongation de la période de collecte des données et un accroissement des coûts. Troisièmement, l'accroissement susmentionné peut accentuer l'ampleur de l'erreur autre que l'erreur d'échantillonnage par rapport à celle de l'erreur d'échantillonnage. En

effet, l'addition de modules supplémentaires au questionnaire peut accroître le risque d'erreur autre que l'erreur d'échantillonnage du fait d'une insuffisance de contrôle des questions ou de la fatigue du répondant. Le risque de telles erreurs peut aussi se trouver aggravé par l'utilisation d'un plus grand nombre d'enquêteurs relativement inexpérimentés, rendue nécessaire par l'accroissement du nombre d'entrevues ou de la durée moyenne d'une entrevue.

F. Quelques mises en garde concernant la détermination des coûts des enquêtes

25. Pour que les comparaisons des coûts se fassent sur des bases raisonnablement identiques entre les enquêtes et d'un pays à un autre, il faut tenir compte d'un certain nombre de facteurs. Premièrement, dans les pays en développement et les pays en transition, les enquêtes sont parrainées par diverses organisations, qui ont souvent des politiques et des procédures comptables différentes. Par exemple, pour certaines de ces organisations, il peut être important de faire la distinction entre les coûts pour l'organisation et le coût global de l'enquête.

26. Deuxièmement, il peut être nécessaire de compter de façon comparable l'aide en nature, par exemple, les véhicules de transport du personnel de terrain. Dans certains cas, cette aide en nature peut être apportée par l'office national de statistique; celui-ci peut en effet affecter des membres de son personnel permanent à une enquête parrainée de source internationale. Bien que ces coûts soient parfois considérés comme aide en nature et exclus du budget détaillé, ils représentent néanmoins un coût d'opportunité dans la mesure où l'enquête est une activité supplémentaire qui prive l'office national de statistique d'une possibilité de consacrer son temps à d'autres travaux.

27. On peut en dire de même de l'offre d'assistance technique extérieure, qui peut être particulièrement importante pour les pays sans infrastructure ni expérience des enquêtes. Dans bien des cas, cette assistance technique est dispensée en nature par les institutions internationales qui effectuent ou parrainent ces enquêtes et ne figure donc pas directement dans le budget de ces dernières. En revanche, parfois, cette assistance technique est sous-traitée et donc incluse dans le budget. Ainsi, l'enquête de type LSMS effectuée en 1998 au Turkménistan a reçu une assistance technique du Research Triangle Institute (RTI) au titre d'un contrat avec la Banque mondiale.

28. Troisièmement, en raison de la structure hiérarchique des coûts (équation 1) donnée à la section B, il importe de faire la distinction entre le coût total d'une enquête et le coût par entrevue. Par exemple, du fait de la disponibilité de ressources plus substantielles et du souci d'obtenir des estimations plus fiables observé à l'échelon sous-national, les grands pays en développement ou en transition tendent à utiliser de plus vastes échantillons (Fonds des Nations Unies pour l'enfance, 2000, chap. 4). En raison des coûts élevés des transports et de la rémunération d'un personnel plus nombreux, le coût total des enquêtes tend à être plus élevé dans les pays plus vastes. En revanche, dans ces pays, le coût par entrevue peut parfois être moins élevé en raison des économies d'échelle et de la répartition des coûts fixes sur un plus grand échantillon.

29. Quatrièmement, certains aspects spéciaux de la conception de l'échantillon compliquent parfois l'évaluation des coûts totaux et des coûts par entrevue. Par exemple, les coûts peuvent se trouver alourdis par le recours au suréchantillonnage ou par l'utilisation d'échantillons sélectionnés pour assurer la réalisation d'objectifs de précision pour certaines sous-catégories de population réduites ou difficiles à identifier à partir des informations du cadre (par exemple, les familles avec des enfants de moins de cinq ans). Enfin, dans les enquêtes sur des populations aux familles de tailles

très variables, il peut être nécessaire de distinguer les coûts par ménage contacté des coûts par entrevue.

G. Résumé et conclusions

30. La plupart des enquêtes effectuées dans les pays en développement et les pays en transition s'inscrivent dans un contexte de graves contraintes budgétaires et d'incertitudes jusque dans l'exécution du budget approuvé. C'est pourquoi l'analyse des facteurs qui influent sur le coût est l'un des aspects les plus importants du processus de conception et de planification des enquêtes dans ces pays. Dans ce chapitre, nous avons présenté un cadre pour cette analyse et nous avons examiné dans quelle mesure les coûts de l'enquête et certains autres éléments connexes sont transposables dans des pays similaires des points de vue de la conception de l'enquête, de la distribution de la population et d'autres facteurs.

31. Nous avons utilisé les grandes enquêtes nationales pour illustrer les questions clés, mais notre analyse peut s'appliquer aux nombreux types d'enquêtes plus modestes effectuées par les systèmes nationaux de statistiques des pays en développement et des pays en transition. Dans la mesure où l'on peut identifier des structures de coût similaires, on peut utiliser les données sur les éléments du coût d'une enquête effectuée dans un pays comme indications pour la conception d'une enquête semblable dans un autre pays, ou pour concevoir plus efficacement une nouvelle enquête dans le même pays. On a souligné la grande disparité des coûts des enquêtes selon l'ampleur de l'infrastructure dont dispose le pays pour les effectuer. On a également insisté sur la nécessité de tenir compte de certaines limites dans les comparaisons des coûts globaux des enquêtes entre divers pays.

32. Pour conclure, nous tenons à rappeler certains points importants concernant le coût des enquêtes dans les pays en développement et les pays en transition :

- a) Bien qu'une analyse minutieuse des éléments de coût révèlent des structures de coût communes selon les pays ou selon les enquêtes, il ne faut pas perdre de vue que le budget est souvent non seulement propre au pays, mais aussi au moment. Il est donc nécessaire de réunir des données de coût et de préparer un rapport administratif décrivant les divers éléments de coût à chaque phase du processus pour chaque enquête sur les ménages. Il faut également analyser les variances et leurs composantes. Ces informations sur les coûts et les variances sont utiles à deux titres : premièrement, pour les grandes décisions de gestion et de budget; et deuxièmement, pour montrer comment différents éléments de coût et de variance influent sur les décisions de conception de l'échantillon. En règle générale, l'analyse des coûts et des variances et de leurs éléments à chaque stade du processus d'enquête devrait faire partie intégrante des méthodes de fonctionnement des offices nationaux de statistique des pays en développement et des pays en transition;
- b) Bien que le budget global de l'enquête comprenne à la fois des coûts fixes et des coûts variables, ce sont ces derniers qu'il faut suivre de près et sur lesquels il faut agir lors de l'établissement de ce budget. Certains coûts fixes, tels que ceux de la coordination de la planification de l'enquête par plusieurs organismes publics et ceux d'une publicité visant les répondants potentiels, ne sont pas du ressort du concepteur de l'enquête et, de toute façon, sont trop fonction du pays, du moment et de l'enquête envisagée;
- c) Comme on le verra au chapitre XIV, il y a une différence concernant les considérations budgétaires selon que l'enquête est payée par les utilisateurs

ou inscrite au budget du pays. Alors que la première est bien conçue et exécutée sans trop de heurts et que tous ses éléments critiques sont payés d'avance, la seconde est généralement sujette aux contraintes budgétaires et aux crédits alloués par le pays. Pour cette dernière, il y a souvent une grande disparité entre le budget prévu et le budget réel, due non pas à des considérations de précision mais de disponibilité des fonds en regard des autres priorités budgétaires du pays;

- d) Étant donné le contexte budgétaire très serré dans lequel s'inscrivent la plupart des enquêtes dans les pays en développement et les pays en transition, il importe que le concepteur examine les moyens non monétaires d'alimenter le budget de l'enquête ou d'exécuter certains de ses éléments sans les inscrire à son budget. Par exemple, il peut être possible de partager l'infrastructure avec une autre enquête, d'utiliser un sous-échantillon d'unités déjà choisies pour une enquête précédente et de faire en sorte que l'enquêteur recueille des données pour plusieurs enquêtes à la fois. Il conviendrait également d'envisager de chiffrer certains aspects d'une enquête d'après le temps nécessaire à leur exécution;
- e) Jusqu'ici, nous avons souligné que le coût d'une enquête peut s'accroître considérablement si le pays ne dispose pas de l'infrastructure nécessaire et d'une capacité générale de statistique. La mise en place et le renforcement de cette infrastructure est donc un bon investissement qui devrait permettre, à long terme, d'alléger le budget des enquêtes dans les pays en développement et les pays en transition. L'un des moyens les plus efficaces pour ces pays de se doter d'une telle infrastructure et de promouvoir le développement général de leurs statistiques est de recourir à la coopération technique entre leurs offices nationaux de statistique et les systèmes statistiques plus développés, en collaboration avec les organisations internationales de statistique et de financement et d'autres parties prenantes. Toutefois, pour donner des résultats positifs pour les pays bénéficiaires, cette coopération technique doit être bien conçue et bien exécutée. Des directives pratiques pour une bonne coopération technique dans le domaine des statistiques ont été ébauchées par les Nations Unies (1998, annexe) et approuvées par la Commission de statistique des Nations Unies lors de sa treizième session, le 4 mars 1999.

REMERCIEMENTS

Je remercie les trois auteurs cités en référence ainsi que les participants à la Réunion du Groupe d'experts sur l'analyse des caractéristiques opérationnelles des enquêtes dans les pays en développement et les pays en transition, tenue en octobre 2002 au siège des Nations Unies à New York, de leurs commentaires très constructifs qui ont permis d'améliorer considérablement la première version de ce chapitre. Toutefois, les opinions exprimées ici n'engagent que moi et ne reflètent pas nécessairement la position de l'Organisation des Nations Unies.

RÉFÉRENCES

- Ajayi, O. O. (2002). Budgeting framework for surveys. Communication personnelle.
- Andersen, R., J. Kasper et M. R. Frankel (1979). *Total Survey Error*. San Francisco, Californie, Jossey-Bass.

- Cochran, W.G. (1977). *Sampling Techniques*, 3^e éd. New York, Wiley.
- de Vries, W. (1999). Are we measuring up? questions on the performance of national statistical systems. *International Statistical Review*, vol. 67, pp. 63-77.
- Fonds des Nations Unies pour l'enfance (2000). *End-Decade Multiple Indicator Cluster Survey Manual*. New York, Fonds des Nations Unies pour l'enfance.
- Grosh, M. E., et J. Muñoz (1996). *A Manual for Planning and Implementing the Living Standards Measurement Study Survey*. Living Standards Measurement Study Working Paper, No. 126. Washington, D.C., Banque internationale pour la reconstruction et le développement, Banque mondiale.
- Groves, R. M. (1989). *Survey Errors and Survey Costs*. New York, Wiley.
- Kish, L. (1965). *Survey Sampling*. New York, Wiley.
- _____ (1976). Optima and proxima in linear sample designs. *Journal of the Royal Statistical Society, Série A*, vol. 139, pp. 80-95.
- Linacre, S. J. et D. J. Trewin (1993). Total survey design: application to a collection of the construction industry. *Journal of Official Statistics*, vol. 9, pp. 611-621.
- Nations Unies (1998). Some guiding principles for good practices in technical co-operation for statistics: note by the Secretariat. E/CN.3/1999/19. 15 octobre 1999.
- Yansaneh, I. S., et J. L. Eltinge (2000). Design effect and cost issues for surveys in developing countries. *Proceedings of the Section on Survey Research Methods*. Alexandria, Virginie, American Statistical Association, pp. 770-775.

Annexe

Cadre budgétaire pour les enquêtes en grappes à indicateurs multiples (MICS) du Fonds des Nations Unies pour l'enfance (UNICEF)

Catégories de coûts	Catégories d'activité						Diffusion et analyse plus approfondie	
	Préparation/ sensibilisation	Enquête pilote	Conception de l'enquête et préparation de l'échantillon	Formation	Exécution de l'enquête principale	Entrée des données		Traitement et analyse des données
Personnel								
Indemnités journalières								
Transports								
Fournitures								
Équipement								
Autres coûts								
COÛTS TOTAUX								
Organismes d'exécution (noms)								

Renseignements supplémentaires

1. Taille de l'échantillon : nombre de ménages : _____ nombre de grappes : _____
2. Durée de l'énumération : nombre de jours : _____
3. Durée de la formation des énumérateurs : nombre de jours : _____
4. Nombres d'énumérateurs/de superviseurs sur le terrain : énumérateurs : _____ superviseurs : _____
5. Entrée des données : caractères frappés par questionnaire : nombre : _____
6. Contribution de l'UNICEF : \$ _____

CADRE DE DÉTERMINATION DES COÛTS

Éléments inclus par catégorie de coût et d'activité

Catégories de coût	Catégories d'activité
Personnel (salaires)	Préparation/sensibilisation
Honoraires de consultant	Préparation du questionnaire
Superviseurs sur le terrain	Préparation de modèles de tableau
Enquêteurs/énumérateurs	Traduction et retraduction
Chauffeurs	Essai préalable du questionnaire
Traducteurs	Publicité avant et après énumération
Guides locaux	
Préposés à l'entrée des données	Enquête pilote
Programmeurs sur ordinateur	Formation
Heures supplémentaires	Collecte de données
Primes d'encouragement	Analyse des données
Comité de coordination	Rapport sur l'enquête pilote
Indemnités journalières (chambre et repas)	Conception de l'enquête et préparation de l'échantillon
Superviseurs sur le terrain	Planification
Enquêteurs/énumérateurs	Préparation de l'échantillon
Chauffeurs	
Traducteurs	Formation
Guides locaux (allocation repas)	Préparation des matériels didactiques
Consultants/observateurs	Traduction dans la langue de formation
	Mise en œuvre de la formation
Transports	Exécution de l'enquête principale
Location de véhicules	Exécution
Allocation pour transports publics	Suivi et supervision
Carburant	Extraction des données
Frais d'entretien	
Visites de consultants	Entrée des données
	Entrée des données
	Repérage des erreurs
Fournitures	Traitement et analyse des données
Fournitures de bureau (papier, crayons, stylos, etc.)	Traitement des données
Cartes d'identité	Correction des données
Enveloppes pour l'envoi des questionnaires	Établissement d'indicateurs
Informatique, fournitures (papier, disquettes, rubans, cartouches)	Tableaux d'analyse
Équipement	Rédaction de rapport
Matériel anthropométrique (balances, toises, etc.)	
Autres coûts	Diffusion et analyse complémentaire
Impression (questionnaire, etc.)	Impression du rapport
Photocopies des cartes, listes et manuels d'instruction	Distribution
Entretien des matériels	Réunions bilans
Communications (téléphone, télécopie, timbres, etc.)	Analyse complémentaire
Contrats (informatique, rédaction de rapports)	

Chapitre XIII

Modèle de détermination des coûts d'une enquête sur les revenus et les dépenses

HANS PETTERSSON

Statistics Sweden
Stockholm, Suède

BOUNTHAVY SISOUPHANTHONG

Centre national de statistiques
Vientiane,
République démocratique populaire lao

RÉSUMÉ

Le présent chapitre décrit le travail d'établissement d'un modèle de détermination des coûts d'une enquête sur les dépenses et la consommation en République démocratique populaire lao. Il commence par une brève description des modèles de détermination des coûts et des problèmes liés à l'estimation des composantes de ce modèle, puis décrit l'Enquête de 2002 sur les dépenses et la consommation de la République démocratique populaire lao. Un modèle de détermination des coûts, conçu sur la base des estimations budgétaires de l'enquête, est utilisé pour calculer les tailles optimales des grappes avec diverses hypothèses concernant les taux de cohérence des grappes. Ce chapitre se termine par une analyse de l'efficacité du modèle d'échantillon choisi par comparaison avec d'autres conditions optimales.

Termes clés : conception de l'enquête, coûts de l'enquête, efficacité, modèle de détermination des coûts, taille optimale de l'échantillon.

A. Introduction

1. La conception d'un échantillon en grappes à plusieurs degrés implique un certain nombre de décisions. Il importe notamment de décider comment allouer au mieux l'échantillon entre les différentes phases. En général, la mise en grappes de l'échantillon a des effets contraires sur les coûts et les variances : il réduit les coûts et accentue les variances. La conception économique d'un échantillon à plusieurs degrés exige du statisticien chargé de l'échantillonnage qu'il estime et équilibre ces influences. Pour ce faire, il lui faut de bonnes informations sur les variances imputables aux différentes phases de l'échantillonnage ainsi que sur les coûts variables qui sont fonction de la taille de l'échantillon à chaque phase.

2. Si des modèles de variance ont été élaborés pour beaucoup de modèles à plusieurs degrés de type courant, l'élaboration de modèles de détermination des coûts a moins retenu l'attention des statisticiens. Aujourd'hui, on dispose de modèles concernant la variance et les effets liés à la conception au moins pour les estimations les plus importantes de nombreuses enquêtes effectuées dans les pays en développement. L'utilisation de modèles de détermination des coûts pour la conception d'échantillons est moins courante. Le problème est dû en partie au peu d'informations détaillées sur

les coûts des enquêtes effectuées par de nombreux instituts nationaux de statistiques, de sorte qu'il est difficile d'établir avec précision le budget d'une enquête et de concevoir un modèle de détermination des coûts réaliste.

3. Dans le présent chapitre, nous examinerons brièvement les modèles de détermination des coûts et indiquerons comment utiliser ces modèles avec les modèles de variance pour déterminer la taille optimale de l'échantillon pour les unités primaires d'échantillonnage d'un modèle à deux phases. Nous présenterons un modèle de détermination des coûts pour une enquête sur les dépenses et la consommation en République démocratique populaire lao et utiliserons ce modèle pour calculer la taille optimale des échantillons à sélectionner au sein des UPE.

B. Modèles de détermination des coûts et estimations des coûts

Modèles de détermination des coûts

4. Un modèle simple de détermination des coûts pour un échantillon à deux phases peut être représenté par l'équation :

$$C = C_0 + C_1 \cdot n + C_2 \cdot n \cdot m \quad (1)$$

où n = le nombre d'unités primaires d'échantillonnage (UPE) de l'échantillon; m = le nombre d'unités secondaires d'échantillonnage (USE) [par exemple, les ménages] de l'échantillon de chaque UPE; C_0 = les coûts fixes d'exécution de l'enquête, indépendamment du nombre d'UPE de l'échantillon et du nombre d'USE par UPE, y compris les coûts de planification et de conception de l'enquête, les coûts des travaux préparatoires, les coûts de gestion de l'enquête et les coûts de traitement et d'analyse des données et de présentation des résultats (certains des coûts de traitement des données sont fonction de la taille de l'échantillon et ne sont donc pas des coûts fixes, mais cela n'est pas pris en compte ici); C_1 = le coût moyen d'addition d'une UPE à l'échantillon, c'est-à-dire les coûts de déplacement des enquêteurs et des superviseurs entre les UPE et leur siège, et d'une UPE à une autre (coût du carburant, salaires des chauffeurs) et les salaires des enquêteurs, ainsi que le coût d'obtention de cartes et d'autres documents pour l'UPE, le coût d'organisation de l'enquête à l'échelon local, ce qui implique notamment des rencontres avec les responsables locaux et l'obtention de leur autorisation, et le coût d'établissement de listes et de sélection d'unités d'habitation/des ménages au sein de l'UPE; C_2 = le coût moyen d'adjonction d'un ménage supplémentaire à l'échantillon, y compris les coûts de localisation de ce ménage et de prise de contact et d'entrevue avec lui, c'est-à-dire les salaires et indemnités journalières de l'enquêteur et du superviseur et les coûts de déplacement de ceux-ci à l'intérieur des UPE.

5. Ce modèle de détermination des coûts est simple, comparé aux modèles plus complexes existants. Hansen, Hurwitz et Madow (1953) ont conçu un modèle qui isolait les coûts de déplacement entre UPE, où :

$$C = C_0 + C_1 \cdot n + C_2 \cdot n \cdot m + C_3 \cdot \sqrt{n} \quad (2)$$

Le coût d'addition d'une UPE (C_3) comprend le coût de déplacement de l'enquêteur entre son lieu d'origine et la première UPE visitée et de retour de la dernière UPE visitée pour la collecte de données à son lieu d'origine, mais pas les coûts de déplacement d'une UPE à une autre, qui sont inclus dans le terme $C_2 \cdot n \cdot m$. Des modèles isolant ces deux types de déplacements ont également été proposés (Kalsbeek, Mendoza et Budescu, 1983). Groves (1989) donne une analyse relativement générale des modèles de détermination des coûts, y compris de diverses formes complexes non linéaire, discontinue, par degrés. Toutefois, la complexité mathématique des modèles de détermination

tion des coûts rend plus difficile la recherche de la forme optimale. De plus, le manque de données précises empêche souvent le recours à des modèles complexes. Dans ce chapitre, nous utiliserons le modèle simple (1) et nous supposerons que les unités de la seconde phase sont les ménages.

Estimations des coûts

6. En général, le responsable de la gestion de l'enquête, se fondant sur l'expérience d'enquêtes précédentes de type similaire, a une idée assez précise du temps nécessaire à l'exécution des diverses opérations de l'enquête. Cette expérience (ou celle d'enquêtes pilotes) lui permet souvent d'établir des estimations raisonnables du temps à consacrer à localiser et à interroger chaque ménage, et par conséquent, de C_2 . En revanche, l'estimation de C_0 , qui implique l'allocation des coûts indirects et des coûts du personnel travaillant à plusieurs projets ou activités, est plus problématique. Il est souvent difficile d'estimer le temps nécessaire au personnel administratif, aux spécialistes et aux superviseurs. Habituellement, les enquêtes précédentes ne donnent pas de bons éléments de jugement sur les coûts de ces types de personnel. De plus, de nombreuses enquêtes utilisent une assistance technique (AT) dispensée par les bailleurs de fonds étrangers, et il peut être difficile de déterminer le temps passé par ces consultants sur une enquête particulière.

7. Le calcul d'une estimation raisonnable de C_1 est souvent difficile car il implique la détermination de l'effet du déplacement d'un enquêteur supplémentaire lorsqu'une UPE est ajoutée à l'échantillon. Ce déplacement dépend de l'étendue de la zone couverte, du nombre d'UPE affectées à chaque enquêteur et du parcours de chaque enquêteur. Il comprend les déplacements d'une UPE à une autre pour la collecte de données et du lieu d'origine au lieu de l'enquête.

8. Il n'y a pas de moyen facile de surmonter les difficultés inhérentes à l'établissement de bonnes estimations des coûts. Une comptabilité analytique précise et assez détaillée des coûts d'enquêtes précédentes ou d'enquêtes pilotes est très utile. Outre l'expérience passée et les opérations pilotes, on peut déterminer les coûts en instaurant un suivi spécial des coûts pour les enquêtes en cours, ce qui se fait notamment pour l'Enquête nationale sur la santé aux États-Unis d'Amérique (*National Health Interview Survey*) [Kalsbeek, Botman et Massey, 1994].

C. Modèles de détermination des coûts pour une conception efficace d'échantillons

9. L'établissement de modèles de détermination des coûts peut servir à deux usages :

- À des fins budgétaires, pour établir le budget d'une enquête sur la base des coûts unitaires du modèle et de la taille prévue des échantillons à divers stades;
- Pour concevoir un échantillon efficace en combinant le modèle de détermination des coûts à un modèle d'erreur d'échantillonnage.

10. Dans ce chapitre, nous nous efforcerons essentiellement d'utiliser les modèles de détermination des coûts pour parvenir à une conception efficace. Nous supposerons un échantillon à deux phases, avec les ménages sélectionnés au sein des UPE dans la seconde phase. Le problème peut être posé en ces termes : étant donné la structure des coûts représentée dans le modèle de détermination des coûts, comment l'échantillon devrait-il être distribué entre les deux phases de l'échantillonnage. Des modèles distincts de détermination des coûts sont généralement établis pour les stra-

tes urbaines et rurales et, dans certains cas, pour d'autres strates. Dans ces cas, il faut aussi prévoir de distribuer l'échantillon entre ces différentes strates.

11. Nous n'avons pas à tenir compte des coûts fixes (C_0) lorsque nous visons une conception efficace; l'élément important est le coût des travaux de terrain : $C_1 \cdot n + C_2 \cdot n \cdot m$. Le coût estimatif des travaux de terrain pour chaque entrevue (C_f) s'obtient en divisant les coûts totaux des travaux de terrain par le nombre d'entrevues ($n \cdot m$), ce qui donne :

$$C_f = C_2 + C_1/m \quad (3)$$

La variance peut s'exprimer ainsi :

$$Var = V \cdot (1 + \rho (m - 1)) \quad (4)$$

où V est la variance pour un simple échantillonnage aléatoire des ménages; ρ est le taux de cohérence (Kish, 1965) [voir aussi chap. VI ci-dessus]; et m est la taille de l'échantillon au sein des UPE.

L'équation (3) montre clairement que l'on peut réduire les coûts des travaux de terrain par entrevue (C_f) en choisissant m aussi grand que possible. L'équation (4) montre tout aussi clairement que la variance augmente à mesure qu'augmente m (et que la variance est minimale lorsque $m = 1$). Le nombre optimum de ménages, m_{opt} , est la valeur de m qui limite le plus possible $Var \cdot C_f$ où :

$$Var \cdot C_f = V \cdot (1 + \rho (m - 1)) \cdot (C_2 + C_1/m) \quad (5)$$

Il a été montré (Kish, 1965) que l'on peut déterminer la taille optimale de l'échantillon en utilisant l'équation :

$$m_{opt} = \sqrt{\frac{C_1 \cdot (1 - \rho)}{C_2 \cdot \rho}} \quad (6)$$

12. Le premier facteur de l'équation (6), C_1/C_2 , est le ratio entre les coûts unitaires de la première et de la deuxième phase. Le coût d'addition d'une UPE à l'échantillon (C_1) est toujours plus élevé que le coût d'addition d'un nouveau ménage à une UPE sélectionnée (C_2), donc, le ratio de coût est toujours nettement supérieur à 1,0. Plus ce ratio est élevé, plus il est coûteux de choisir une nouvelle UPE par comparaison avec la sélection d'un plus grand nombre de ménages dans les UPE sélectionnées; par conséquent, nous devrions choisir de préférence d'ajouter des ménages dans les UPE déjà sélectionnées.

13. La grandeur ρ mesure la cohérence interne de l'UPE. Lorsque cette cohérence est élevée, il est préférable de ne pas prendre un gros échantillon de ménages au sein des UPE, dans la mesure où le gain en informations obtenues de chaque nouveau ménage de l'échantillon est modeste (les ménages étant très semblables). Cela se reflète dans le second facteur de l'équation (6). Lorsque ρ est élevé, ce facteur et m_{opt} deviennent modestes (pour un ratio de coûts donné).

14. Les valeurs de ρ découlent souvent des estimations des effets liés à la conception provenant d'enquêtes précédentes. Pour beaucoup de variables démographiques, ces valeurs tendent à être faibles, souvent inférieures à 0,01. Pour beaucoup de variables socio-économiques, elles peuvent être supérieures à 0,1 et, dans certains cas, atteindre 0,2 ou 0,3.

15. Le ratio des coûts doit aussi être déterminé à partir d'enquêtes précédentes. Il convient de souligner qu'il n'est pas nécessaire d'exprimer ce ratio en termes de coûts. Le temps (exprimé en nombre de jours d'enquêteur nécessaires) est souvent utilisé comme unité au lieu des coûts : le calcul est approximativement le même (certains coûts de déplacement peuvent être ignorés). Le niveau du ratio des coûts dépend

de la conception du travail de terrain. Pour une enquête où le temps consacré à une entrevue est très bref, ce ratio peut être de 20-50. Par exemple, si le temps nécessaire par UPE est de trois jours et si l'enquêteur peut couvrir 10 ménages par jour, le ratio des coûts (calculé comme ratio de temps T_1/T_2) est de 30 ($T_1 = 3$ jours et $T_2 = 0,1$ jour). Pour les enquêtes nécessitant de très longues entrevues, le ratio des coûts peut être inférieur à 10.

16. Ces calculs peuvent donner l'impression que l'on peut obtenir une réponse précise à la question de savoir combien de ménages choisir dans chaque UPE. Tel n'est presque jamais le cas, cependant, pour plusieurs raisons, à savoir :

- Le modèle de détermination des coûts est une approximation assez grossière de la réalité. On a besoin de simplification pour que le modèle soit utilisable (comme indiqué à la section B).
- Les estimations des coûts et les valeurs de ρ sont empreintes d'incertitude.
- L'optimum n'est valable que pour une variable de l'enquête. Si des variables importantes ont différents niveaux de ρ , il n'y a pas une taille optimale unique pour les grappes mais plutôt un certain nombre de tailles optimales différentes.

17. Les calculs donnent une indication assez approximative de la taille optimale de l'échantillon pour différentes valeurs de ρ . Cette information peut servir à décider d'une taille de l'échantillon qui convienne assez bien pour toutes les variables importantes, mais d'autres facteurs, souvent liés à des contraintes pratiques concernant le travail de terrain, sont à prendre en compte lors de la décision finale.

D. Étude de cas : l'Enquête de 2002 sur les dépenses et la consommation en République démocratique populaire lao

18. Le Centre national de statistique (CNS) de la République démocratique populaire lao a effectué trois enquêtes sur les dépenses et la consommation au cours de la dernière décennie. La première (LECS-1) a eu lieu en 1992-1993, la deuxième (LECS-2) en 1997-1998 et la troisième (LECS-3) en 2002-2003. La présente section donne une description de LECS-3.

19. Les données de ces enquêtes sont destinées à un certain nombre d'usages, dont le plus important est de produire des estimations nationales de la consommation et de la production des ménages pour la comptabilité nationale. Ces estimations portent notamment sur la production agricole et les activités commerciales des ménages.

Conception de l'échantillon pour LECS-3

20. L'échantillon était constitué de 8 100 ménages sélectionnés par un échantillonnage à deux phases. Les villages ont servi d'unités primaires d'échantillonnage (UPE). Ils ont été stratifiés en 18 provinces et, à l'intérieur de celles-ci, en zones urbaines ou rurales. Les villages ruraux ont à nouveau été stratifiés en villages « avec » ou « sans » route d'accès. L'échantillon total de première phase comprenait 540 villages. Il a été attribué aux provinces proportionnellement à la racine carrée de leur population telle qu'elle ressortait du recensement. Dans chaque province, les UPE ont été sélectionnées selon une procédure de probabilité systématique proportionnelle à la population.

21. La liste des ménages des villages sélectionnés a été établie avant l'enquête. Dans chaque village, 15 ménages ont été sélectionnés par échantillonnage systéma-

tique, ce qui a produit un échantillon de 8 100 ménages. La décision de sélectionner 15 ménages par village était fondée principalement sur des considérations pratiques. À la section E, nous comparerons l'efficacité d'échantillons de 15 ménages aux tailles optimales d'échantillon correspondant à différentes hypothèses concernant les taux de cohérence.

Collecte des données pour LECS-3

22. Les données ont été recueillies au moyen : a) de questionnaires à remplir par les ménages; b) de questionnaires à remplir par les villages; et c) d'un questionnaire sur les prix. Les deux derniers questionnaires ont servi essentiellement de moyens de recueillir des informations supplémentaires pour l'enquête sur les ménages.

23. Une grande partie du questionnaire à remplir par les ménages est restée la même que pour les enquêtes précédentes, à l'exception des modifications apportées aux questions auxquelles il n'avait pas été répondu précédemment de façon satisfaisante. Les données sur les dépenses et la consommation ont été recueillies pour un mois entier sur la base de l'enregistrement journalier des dépenses. À la fin du mois, le ménage a été interrogé sur ses achats de biens durables au cours des 12 mois précédents. Pendant le mois, chaque membre de la famille devait avoir enregistré son emploi du temps sur une période de 24 heures. La consommation de riz « de la veille » a été mesurée pour chaque membre de la famille pour donner une mesure plus précise de la consommation de chacun à chaque repas.

24. Le questionnaire à remplir par le village, qui a été confié au chef du village, portait sur certains points tels que les routes et les transports, l'eau, l'électricité, les équipements de santé, les marchés locaux, les écoles, etc. Le questionnaire sur les prix a été utilisé par les enquêteurs pour recueillir des données sur les prix locaux de 121 produits.

Travail de terrain

25. La mesure de la consommation journalière par la tenue d'un journal par les ménages a imposé une lourde charge non seulement aux ménages, mais aussi aux enquêteurs. De nombreux ménages, surtout dans les zones rurales, ont eu souvent besoin d'aide pour tenir ce journal. Pour garantir une qualité acceptable des données, il a été jugé nécessaire de maintenir les enquêteurs durant tout le mois dans le village plutôt que de leur demander de s'y rendre plusieurs fois pour une série de secondes entrevues ou pour un suivi. Cette décision était également jugée nécessaire par le fait que beaucoup de villages, surtout dans les zones montagneuses, étaient d'accès difficile (certains ne pouvaient être atteints qu'après plusieurs jours de marche).

26. Lors des enquêtes précédentes, des équipes de deux enquêteurs avaient effectué le travail de terrain dans chaque village. Pour LECS-3, un seul enquêteur avait été pressenti. Toutefois, en dernière analyse, certains facteurs liés à la sécurité et le bien-être des enquêteurs ont fait pencher la balance en faveur d'équipes de deux enquêteurs. Ceux-ci ont rendu plusieurs visites aux ménages sélectionnés pendant la période de quatre semaines. Ils ont aussi travaillé avec les chefs des villages pour les aider à remplir le questionnaire destiné au village et pour mettre à jour les registres villageois. Pendant le mois, les enquêteurs ont également recueilli des données sur les prix au marché local.

27. Le personnel de terrain était composé de 180 enquêteurs organisés en 90 équipes de deux. Trente-six superviseurs des offices provinciaux de statistiques et 10 superviseurs du siège central ont supervisé ces équipes.

E. Modèle de détermination des coûts pour le travail de terrain lors de l'Enquête de 2002 sur les dépenses et la consommation (LECS-3)

Estimations des coûts

28. Dans une large mesure, LECS-3 ressemblait aux deux LECS précédentes. L'expérience concernant le temps nécessaire au travail de terrain lors de ces deux enquêtes précédentes a donc permis d'estimer les coûts du travail de terrain de LECS-3.

29. Le tableau XIII.1 présente des estimations du temps nécessaire au travail de terrain pour LECS-3. Des estimations ont été établies respectivement pour les zones urbaines et rurales.

Tableau XIII.1.

Durée estimative du travail de terrain dans un village

	Présentation de l'enquête, établissement de la liste des ménages et sélection de ménages dans les villages, collecte d'informations sur le village		
	Déplacement		Entrevue avec les ménages
	Nombre de jours/village	Nombre de jours/village	Nombre de jours/village
Quartiers urbains (100)			
Superviseurs provinciaux	1,5	0,5	3
Enquêteurs (équipes de deux)	3	7	47
Villages ruraux (440)			
Superviseurs provinciaux	3	0,5	3
Enquêteurs (équipes de deux)	6	7	47

30. Le tableau XIII.2 présente des estimations des coûts du travail de terrain calculées sur la base des estimations de temps du tableau XIII.1. Ces coûts comprennent les coûts de déplacement (généralement par voiture ou par bus) et les indemnités journalières pour le temps de travail sur le terrain. Les personnes travaillant à cette enquête appartenaient sans exception au personnel permanent du CNS et avaient été affectées à l'enquête dans le cadre de leurs activités ordinaires. Par conséquent, les éléments de coût ne comprennent pas leurs salaires.

Tableau XIII.2.

Estimation des coûts de LECS-3 (en dollars des États-Unis par jour)

	Présentation de l'enquête; établissement des listes et sélection des ménages; collecte d'informations dans les villages		
	Coûts de déplacement (indemnités journalières pour la durée du séjour sur le terrain et estimation des coûts de transport)		Entrevues avec les ménages
	A	B	C
Quartiers urbains (100)			
Superviseurs provinciaux	1 540	450	2 710
Enquêteurs (équipes de deux)	2 490	5 060	33 970
Villages ruraux (440)			
Superviseurs provinciaux	15 850	1 990	11 950
Enquêteurs (équipes de deux)	25 560	22 260	149 460
Total	45 440	29 760	198 090

Modèle de détermination des coûts

31. Les colonnes **A** et **B** du tableau XIII.2 indiquent les coûts liés à la sélection et à la préparation des villages pour l'enquête. La somme des éléments figurant dans ces colonnes divisée par le nombre de villages représente le coût moyen (C_1) en dollars des États-Unis d'inclusion d'un village dans l'enquête; pour les quartiers urbains, $C_1 = (1\,540 + 2\,490 + 450 + 5\,060)/100 = 95$; et pour les zones rurales: $C_1 = (15\,850 + 25\,560 + 1\,990 + 22\,260)/440 = 149$. Tous les déplacements sont considérés comme déplacements entre deux quartiers urbains ou deux villages ruraux; par conséquent, tous les coûts de déplacement sont inclus dans C_1 .

32. La colonne **C** du tableau XIII.2 présente les coûts de l'enquête liés aux entrevues avec les ménages. Le principal élément est le temps de l'enquêteur. La somme des éléments de cette colonne représente le coût moyen (C_2), en dollars des États-Unis, d'inclusion d'un ménage dans l'enquête; pour les quartiers urbains, $C_2 = (2\,710 + 33\,970)/(100,15) = 24$; et pour les zones rurales, $C_2 = (11\,950 + 149\,460)/(440,15) = 24$. Lorsque l'on ajoute les valeurs estimatives de C_1 et C_2 , la fonction donnant le coût est la suivante :

$$\text{Quartiers urbains : } C_{\text{fieldwork}} = 95 \cdot n + 24 \cdot n \cdot m \quad (7)$$

$$\text{Zones rurales : } C_{\text{fieldwork}} = 149 \cdot n + 24 \cdot n \cdot m \quad (8)$$

33. Le fait de ne pas avoir inclus les coûts des salaires du personnel permanent aux coûts du personnel entraîne une sous-estimation de C_1 et C_2 et, par conséquent, des coûts du travail de terrain $C_{\text{fieldwork}}$. Toutefois, le plus important pour l'optimisation de la conception est le ratio des coûts C_1/C_2 . Nous estimons que ce ratio n'est affecté que légèrement par l'omission de ces salaires, car celle-ci aurait des effets assez semblables sur C_1 et C_2 .

34. Le ratio des coûts entre les échantillons des première et seconde phases est $C_1/C_2 = 95/24 = 3,9$ pour les zones urbaines et $149/24 = 6,1$ pour les zones rurales. Ces ratios sont relativement faibles, ce qui s'explique par le fait que l'enquête a nécessité un temps considérable d'entrevue et de suivi par ménage pendant le mois passé sur place par l'enquêteur à aider à la tenue d'un journal. À ce titre, LECS-3 a été une enquête assez inhabituelle.

Taille optimale de l'échantillon pour les villages

35. Lors des deux précédentes LECS, les équipes de deux enquêteurs avaient une charge de travail de 20 ménages dans chaque village. Pour LECS-3, la taille de l'échantillon a été réduite à 15 ménages, car les entrevues ont été beaucoup plus longues que lors des deux enquêtes précédentes et LECS-3 comprenait un questionnaire sur les prix qui ne figurait pas dans ces enquêtes.

36. Avec deux enquêteurs dans le village, couvrant un échantillon de 15 ménages, quelle a été l'efficacité de la conception de l'enquête ? Le modèle de détermination des coûts et le modèle de variance ont pu être utilisés pour évaluer l'efficacité relative d'échantillons de 15 ménages.

37. Le tableau XIII.3 indique la valeur optimale de m pour différentes valeurs de ρ . L'efficacité relative de notre modèle est présentée aux rangées trois et quatre. Elle se calcule comme le ratio entre le minimum de $Var.C_f$ [voir (5)] et la valeur réelle de $Var.C_f$ pour une valeur donnée de ρ et un échantillon de 15. L'efficacité est raisonnablement élevée pour des valeurs ρ allant jusqu'à 0,10 et plutôt faible et tendant à se détériorer lorsque ρ atteint 0,2 et plus.

Tableau XIII.3
Tailles optimales de l'échantillon (m_{opt}) et efficacité relative de la conception ($m = 15$) pour différentes valeurs de ρ

	$\rho = 0,01$	$\rho = 0,05$	$\rho = 0,10$	$\rho = 0,15$	$\rho = 0,2$	$\rho = 0,25$
m_{opt} , quartiers urbains	20	9	6	5	4	4
m_{opt} , zones rurales	24	11	8	6	5	4
Efficacité relative (pourcentage) quartiers urbains	99	94	82	73	66	61
Efficacité relative (pourcentage) zones rurales	96	98	89	81	75	70

38. Lors des précédentes LECS, les calculs de ρ avaient montré de nettes différences entre zones urbaines et rurales pour les variables importantes. Les valeurs de ρ étaient nettement plus faibles dans les zones urbaines que dans les zones rurales. Elles se situaient dans une fourchette de 0,04 à 0,08 pour de nombreuses estimations urbaines, ce qui correspondait à un échantillon optimal de huit ou neuf ménages. Notre échantillon de 15 ménages par UPE présente une efficacité relative de 85 à 95 %. Dans les zones rurales, ρ se situe dans une fourchette de 0,11 à 0,20, ce qui correspond à un échantillon optimal de cinq à sept ménages. Notre échantillon affiche une efficacité relative de 75 à 88 %. Un certain degré d'incertitude demeure, surtout quant aux valeurs de ρ auxquelles nous devons nous attendre pour certaines variables importantes de LECS-3. Néanmoins, nous pouvons conclure sans danger que notre échantillon de 15 ménages est supérieur au nombre optimal.

39. Quelles sont les conclusions pratiques que l'on peut tirer de ces résultats pour les LECS à venir ? Les pertes d'efficacité sont faibles dans les zones urbaines; par conséquent, nous pouvons décider de nous en tenir à des échantillons de 15 ménages. En revanche, nous souhaiterions réduire la taille des échantillons par UPE dans les zones rurales. Cependant, du fait de l'organisation actuelle du travail de terrain, qui prévoit un séjour d'un mois complet des enquêteurs dans les UPE, il est difficile de réduire sensiblement la charge de travail. Cela veut dire que les enquêteurs ne seront pas pleinement occupés pendant ce séjour. On peut envisager de leur confier d'autres tâches, par exemple d'effectuer des études des communautés de la région, mais la viabilité de cette option n'a pas été pleinement examinée.

F. Conclusions

40. Un modèle de détermination des coûts, conçu pour LECS-3 et analysé, montre que le ratio des coûts C_1/C_2 de l'enquête était assez faible. La principale raison est l'appui apporté par les enquêteurs à la tenue du journal par les ménages, qui immobilise les enquêteurs pendant un mois dans les villages. À cet égard, LECS-3 a été une enquête d'un genre assez inhabituel, comparée aux autres enquêtes sur le revenu et les dépenses des ménages, pour lesquelles le temps consacré aux entrevues était généralement plus court.

41. Les calculs des tailles optimales des échantillons au sein des UPE montrent que l'échantillon actuel de 15 ménages est supérieur au nombre optimal, surtout dans les zones rurales. Néanmoins, du fait de certaines contraintes pratiques, il pourrait être difficile de réduire ce nombre.

42. Il convient de signaler que le modèle de détermination des coûts n'est qu'une approximation grossière de la réalité, dont aucun modèle simple ne peut saisir toute la complexité. On pourrait construire des modèles plus complexes, y compris diverses expressions des coûts par degrés. Toutefois, la complexité de l'expression ma-

thématique des modèles de détermination des coûts complique souvent la détermination des conditions optimales.

RÉFÉRENCES

- Groves, R. M. (1989). *Survey Error and Survey Costs*. New York, John Wiley and Sons.
- Hansen, M. H., W. N. Hurwitz et W. G. Madow (1953). *Sample Survey Methods and Theory*, vol. I. New York, John Wiley and Sons.
- Kalsbeek, W., O. M. Mendoza et D. V. Budescu (1983). Cost models for optimum allocation in multi-stage sampling. *Survey Methodology*, vol. 9, n° 2, pp. 154-177.
- Kalsbeek W. D., S. L. Botman et J. T. Massey (1994). Cost efficiency and the number of allowable call attempts in the National Health Interview Survey. *Journal of Official Statistics*, vol. 10, n° 2, pp. 133-153.
- Kish, L. (1965). *Survey Sampling*. New York, John Wiley and Sons.

Chapitre XIV

Élaboration d'un cadre pour l'établissement du budget des enquêtes sur les ménages dans les pays en développement

ERICA KEOGH

Statistics Department, University of Zimbabwe
Harare, Zimbabwe

RÉSUMÉ

Le présent chapitre vise à faire des recommandations pour l'établissement d'un budget prudent et logique pour une enquête. Il montre aux lecteurs qu'il y a deux façons de concevoir un tel budget, par poste comptable ou par activité prévue, et il les encourage à élaborer leur budget en utilisant l'approche consistant à décrire les postes comptables pour chaque activité de l'enquête. Le produit fini est une matrice des coûts qui peut aussi être utilisée pour enregistrer les dépenses réelles tout au long de l'enquête. Il est fortement recommandé de documenter et d'analyser les coûts réels de l'enquête, afin d'offrir des éléments de jugement pour les enquêtes à venir. Ce chapitre met constamment l'accent sur la relation critique entre la conception d'une enquête par sondage et l'établissement de son budget.

Termes clés : conception de l'enquête, budgets de l'enquête, exécution de l'enquête.

A. Introduction

1. Les enquêtes étant coûteuses en temps et en argent, il est essentiel d'établir un budget détaillé des dépenses prévues du début à la fin de l'opération. Il faut également des provisions pour imprévus, pour urgences et pour changements économiques inattendus, de manière que ces événements soient couverts par le budget proposé. Pour ce faire, il faut que le processus budgétaire puisse s'ajuster en fonction de l'ampleur de l'enquête, notamment de la taille des échantillons, afin de pouvoir mieux faire face aux changements imprévus de la situation économique qui risquent de compromettre l'exécution de l'enquête. Le budget de l'enquête doit être considéré comme un élément dynamique du processus, susceptible d'adaptations aux besoins réels en cours de route. Les outils nécessaires pour suivre les dépenses seront conçus parallèlement au budget, et constamment actualisés en fonction des progrès du processus budgétaire.

2. L'ampleur du budget et l'allocation des crédits aux diverses composantes de l'opération ayant un impact direct sur la qualité de l'enquête, on ne saurait trop souvent insister sur l'importance d'une planification et d'un budget détaillés. Nous avons présenté un examen détaillé des coûts de conception des enquêtes sur les ménages au chapitre XII. L'étude des Nations Unies (1984) souligne combien il importe de trouver le juste équilibre entre les coûts et la qualité de la façon suivante : « En théorie, les prio-

rités devraient être déterminées par une analyse des coûts et des avantages de diverses façons possibles d'utiliser les maigres ressources disponibles » (par. 1.5). Souvent, le budget de l'enquête est fixe, et l'on attend du concepteur qu'il conçoive un échantillon avec des niveaux d'erreur acceptables dans les limites de ce budget.

3. L'établissement d'un budget détaillé pour une enquête proposée est souvent un exercice laborieux, qui implique une planification et une préparation minutieuses. De surcroît, les planificateurs se trouvent dans une situation difficile car ils ne peuvent estimer correctement le budget avant qu'un plan définitif ne soit arrêté et ils doivent cependant établir ce budget avant la planification et la conception de l'enquête. C'est ici que l'expérience de l'estimation du budget et des coûts d'enquêtes précédentes joue un rôle important. Il ne faut pas non plus perdre de vue que l'on ne peut envisager une allocation optimale de l'échantillon sans tenir compte des coûts : par exemple, en cas d'échantillonnage stratifié, on peut choisir entre limiter le coût pour un niveau fixé de précision ou obtenir une précision optimale pour des coûts fixes (Scheaffer, Mendenhall et Ott, 1990). Or, souvent les modèles de détermination des coûts ne sont pas réalistes, ne tiennent pas compte de la possibilité de changement de situation pendant l'exécution de l'enquête et ne considèrent généralement les erreurs que sur une variable. Il est donc important de tenir une comptabilité détaillée du budget et des dépenses, afin d'encourager les organismes d'enquête à diffuser leurs informations sur leurs coûts de manière à faciliter la planification des enquêtes à venir.

4. Généralement, les données des enquêtes sont nécessaires aux fins de la planification et/ou des décisions de politique générale; il faut donc pouvoir en disposer le plus tôt possible. Souvent, l'enquête doit s'inscrire dans un calendrier très strict, et des échéances sont spécifiées pour ses divers stades par les organismes de financement. Cependant, l'utilisation d'un petit délai supplémentaire permet parfois l'acquisition de données de bien meilleure qualité; les concepteurs de l'enquête doivent donc être prêts à plaider pour un tel délai lors de l'établissement du budget. En effet, l'allocation d'un temps et/ou d'un budget insuffisant au traitement et à l'analyse des données risque souvent de compromettre la qualité des résultats de l'enquête. Il est donc nécessaire, au stade de l'établissement du budget, de jongler entre le temps, le coût et les erreurs pour arriver au cadre d'opération le plus approprié.

5. Le présent chapitre vise à éclairer le lecteur sur :

- La façon de préparer le budget,
- Les pièges auxquels s'attendre lors de l'exécution de l'enquête,
- Les outils à utiliser pour gérer les finances de l'enquête et rendre compte de sa gestion,

notamment pour ce qui est des enquêtes sur les ménages dans les pays en développement.

B. Considérations préliminaires

1. Les phases de l'enquête

6. Pour commencer, avant d'examiner de façon quelque peu détaillée les principales composantes du budget d'une enquête sur les ménages, il est bon de rappeler les principales phases de l'enquête, car les coûts de chacune de ces phases doivent être planifiés et respectés dans toute la mesure du possible. On peut résumer comme suit les différentes phases de l'enquête :

- Conception et préparation de l'enquête;
- Exécution de l'enquête;
- Établissement d'un rapport sur l'enquête.

Les différentes composantes de ces phases ont été analysées de façon détaillée dans les chapitres précédents.

2. Calendrier de l'enquête

7. Le deuxième élément essentiel à considérer lors de l'établissement du budget est le calendrier d'ensemble de l'enquête. En général, quand on prépare une enquête, des crédits ont été promis sur la base d'une date d'achèvement et, éventuellement, de diverses autres échéances imposées. Pour que l'enquête se déroule sans heurts, il est indispensable d'établir un calendrier réaliste ainsi qu'un cadre budgétaire et de s'y tenir pendant son exécution.

Exemple 1

8. Supposons que l'on ait été chargé d'effectuer une enquête dans une grande ville pour fournir certaines informations de base sur les entreprises du secteur informel, leur fonctionnement et leur succès. Plusieurs donateurs sont intéressés par les résultats car ils désirent apporter une aide sous la forme d'une formation à la gestion et d'un microfinancement aux entrepreneurs méritants. En particulier, ils souhaiteraient s'assurer qu'il est tenu compte des questions d'égalité entre les sexes, et ils aimeraient pouvoir suivre à l'avenir l'impact de leur aide. Ils sont donc prêts à allouer des fonds pour une petite enquête afin d'interroger 500 ménages ou propriétaires de petites entreprises de la ville. Un délai de trois mois sera alloué à la collecte des données, et un mois supplémentaire sera prévu pour l'établissement d'un projet de rapport. Un budget doit être proposé pour cette enquête.

Tableau XIV.1

Proposition de calendrier pour une enquête sur le secteur informel

Tâche	Semaine numéro																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
Consultations avec les donateurs/publicité																		
Conception et expérimentation du questionnaire																		
Conception et sélection de l'échantillon																		
Conception de l'entrée des données																		
Planification de l'analyse des données																		
Recrutement du personnel de terrain																		
Formation des énumérateurs et pilote																		
Impression des questionnaires																		
Travail de terrain et contrôle																		
Entrée et validation des données																		
Correction des données et analyse																		
Production de graphiques et de tableaux																		
Rédaction du rapport																		
Archivage																		

9. Un premier projet de calendrier (tableau de Gantt) est présenté ci-dessus. Lorsque l'on considère le temps disponible pour certaines tâches particulières, on doit estimer le personnel nécessaire pour exécuter ces tâches dans le délai imparti; par exemple, si l'on a prévu quatre semaines pour effectuer 500 entrevues, y compris les visites de rappel, il faut effectuer environ 24 entrevues par jour. Le nombre d'agents de

terrain à prévoir sera fonction de la longueur du questionnaire, du nombre d'entrevues par jour et des distances à parcourir pour se rendre du logement d'un répondant à celui du suivant.

10. Le tableau ci-dessus montre :

- Comment les phases de l'enquête se chevauchent; par exemple, la conception de l'entrée des données se fait en même temps que la conception du questionnaire; l'entrée des données proprement dite commence peu après que les premiers questionnaires sont prêts, et la correction des données peut commencer avant même que toutes les données soient entrées;
- Comment certaines tâches se poursuivent sur toute la durée de l'enquête; par exemple, la préparation du rapport doit être une tâche constante pour les coordonnateurs de l'enquête, car il faut rendre compte de chaque phase de l'étude;
- Comment, dans certains cas, on ne peut commencer une phase avant d'avoir terminé une autre; par exemple, l'impression des questionnaires ne peut se faire qu'une fois terminée l'expérience pilote et là, le créneau prévu pour l'impression est très bref et se situe au même moment que la formation (il est toujours recommandé d'entreprendre les entrevues le plus tôt possible après la formation).

3. Type d'enquête

11. Le budget est fonction du type d'enquête. À cet égard, on distingue deux principaux types d'enquêtes : celles qui sont financées sur le budget de l'État et celles qui sont payées par les utilisateurs.

Enquêtes financées sur le budget de l'État

12. Chaque pays a ses propres institutions chargées d'effectuer des enquêtes périodiques, telles que les enquêtes sur la santé et la nutrition, les enquêtes démographiques sur les ménages, les enquêtes sur les revenus, la consommation et les dépenses et les enquêtes sur l'agriculture et l'élevage. La plupart de ces enquêtes disposent généralement, entre autres :

- D'une infrastructure commune utilisée pour chacune d'entre elles, de sorte qu'elles s'inscrivent dans le cadre d'un programme « intégré »;
- D'un budget financé par le gouvernement central, parfois avec un financement supplémentaire de donateurs;
- D'un personnel permanent pour leur exécution;
- D'un matériel informatique et de moyens de transport.

Autrement dit, les enquêtes font partie des activités courantes de certaines administrations publiques et, à ce titre, s'appuient sur l'expérience des enquêtes précédentes pour l'établissement de leur budget. Elles utilisent généralement un échantillon national représentatif et se fondent souvent sur un calendrier plus ou moins souple, avec des échéances exprimées en mois plutôt qu'en jours. Certains éléments de leur budget présentés ci-dessous dans ce chapitre ne s'appliquent pas nécessairement à certaines d'entre elles.

Enquêtes payées par les utilisateurs

13. Les enquêtes payées par les utilisateurs ne sont pas liées à un programme du gouvernement central mais sont exécutées par des organismes privés financés par

diverses organisations non gouvernementales et divers donateurs, nationaux ou internationaux. Il peut s'agir d'opérations nécessitant rapidement de bons résultats, ou d'opérations devant servir à suivre des programmes et réclamant parfois une analyse approfondie des données pour concevoir des modèles pour des activités futures. Les organismes effectuant ces enquêtes peuvent avoir :

- une infrastructure limitée sur laquelle s'appuyer pour l'enquête;
- une réserve de personnel à laquelle faire appel pour ces études;
- un matériel informatique et des moyens de transport limités;
- des ressources fixes limitées.

ou être bien organisés et avoir effectué récemment un certain nombre d'enquêtes de ce genre. Les ressources fixes et les frais généraux doivent être inscrits au budget, et si l'organisme est privé, il lui faut tenir compte des considérations de profit. En général, la taille des échantillons nécessaires à ces enquêtes n'est pas très grande, et souvent, l'enquête ne se concentre que sur quelques régions géographiques du pays. Souvent, elle est astreinte à des échéances très strictes, malheureusement, au détriment de la qualité des données, en raison d'une planification insuffisamment réaliste.

4. Budgets et dépenses

14. Le budget s'élabore bien avant que commence l'exécution de l'enquête. Le cadre budgétaire doit être conçu et soumis aux organismes de financement avant la planification. Il faut donc formuler certaines hypothèses de base sur la conception de l'enquête lors de l'élaboration du budget. En revanche, les dépenses effectives sont fonction de ce qui se passe réellement lors de l'exécution de l'enquête. Les personnes qui sont chargées de cette exécution doivent être conscientes de cette distinction et se rendre compte que le budget doit couvrir les dépenses encourues. Ces dépenses sont fortement tributaires de phénomènes tels que l'inflation, les variations des taux de change, etc. et, bien entendu, varient d'un pays à l'autre, et parfois, très sensiblement. Il est donc recommandé que le budget s'établisse en fonction d'un nombre d'hommes-jours, des distances à parcourir, ainsi que de prévisions des coûts (exprimés en devises), afin de mieux faire face aux risques d'inflation galopante et d'autres variations inattendues des conditions macroéconomiques du pays. Comme on l'a indiqué précédemment, le budget d'une enquête est un phénomène dynamique, qui nécessite de constantes actualisations, une fois que les dépenses d'exécution se matérialisent.

5. Études précédentes

15. L'expérience passée est toujours riche d'enseignements. Toutefois, l'établissement du budget d'une enquête est une tâche beaucoup plus compliquée que ce à quoi on pourrait s'attendre. Dans le monde entier, on constate une tendance à rendre compte de façon très insuffisante, voire incompétente, des coûts d'une enquête, de sorte qu'il est très difficile d'utiliser les informations passées pour préparer une prochaine enquête. Lorsque l'on demande des informations sur les coûts à une organisation qui a effectué récemment une enquête, l'auteur a constaté que, le plus souvent, on ne pouvait obtenir que les budgets initiaux, et que les allocations s'en écartaient pour différentes raisons, telles que l'inflation. Les coûts réels ne semblaient pas avoir fait l'objet d'un rapport, et toutes les parties semblaient avoir accepté ce fait comme normal, tant que l'opération restait dans les limites du budget. L'examen des budgets d'enquêtes passées a révélé un autre problème, à savoir l'absence d'indication des coûts occultes, tels que ceux correspondant à l'utilisation gratuite de véhicules, le salaire du

directeur, etc. Le fait que ces coûts sont souvent traités comme frais généraux et n'entrent pas dans le budget de l'enquête est donc trompeur pour le chercheur.

16. On espère que la lecture de ce chapitre encouragera les personnes chargées de l'exécution d'enquêtes à suivre les coûts au jour le jour et à en rendre compte pleinement, afin que les chercheurs puissent tirer les leçons du passé. Une documentation complète permettant d'établir le coût d'une enquête par entrevue est très précieuse pour ceux qui cherchent à déterminer le budget d'exercices semblables à l'avenir. Le coût par entrevue doit tenir compte de tous les aspects des coûts réels de l'enquête, y compris la conception, le travail de terrain, le traitement des données et l'établissement d'un rapport, et offre une vision d'ensemble des coûts réels.

C. Principaux postes comptables du cadre budgétaire

17. Il y a deux façons de visualiser le budget et, en fin de compte, les dépenses d'une enquête, à savoir selon les activités et selon les procédures comptables courantes. Lors de l'établissement du cadre budgétaire, il est recommandé de considérer les postes comptables séparément pour chaque activité de l'enquête. On peut ensuite récapituler globalement les postes comptables en s'appuyant sur les informations dont on dispose sur chaque activité, et les réunir pour les présenter aux organismes de financement. En même temps, il serait utile de montrer à ces organismes le budget détaillé de chaque activité, afin d'indiquer les besoins propres à chacune d'entre elles. Le tableau XIV.2 ci-dessous offre un exemple de cette approche utilisant une matrice pour illustrer la nécessité d'un budget depuis chacun de ces deux points de vue.

Tableau XIV.2.

Matrice des postes comptables en regard des activités de l'enquête

	Consultations	Conception	Échantillonnage	Travail de terrain	Traitement des données	Rapport	Total
Personnel							
Transport							
Équipement							
Fournitures							
Autres							
Total							

18. En comparant cette matrice au calendrier du tableau de Gantt présenté au tableau XIV.1, on remarque que le tableau XIV.2 mentionne les mêmes activités que le tableau précédent. Bien que, pour des raisons de place, on ait procédé ici à certains groupements, il faudra indiquer de façon détaillée d'où exactement proviennent les coûts.

19. La présente section vise à définir les postes comptables applicables au budget de l'enquête, tandis que la section D se concentrera sur l'établissement du budget des activités de l'enquête et que la section E fera la synthèse de l'ensemble. Les catégories mentionnées ci-après ne sont pas exhaustives, et l'on pourrait avoir à définir de nouvelles activités propres à l'enquête.

1. Personnel

20. Les traitements et salaires de l'ensemble du personnel doivent être calculés avec soin et incorporés au budget. Les coûts additionnels à prendre en compte ici

comprennent ceux qui pourraient apparaître si l'enquête se prolongeait; par exemple, l'inflation peut nécessiter une hausse des salaires. Il faut aussi prévoir l'éventualité de maladies et la mobilité du personnel.

21. Les salaires versés au personnel doivent être conformes aux conditions locales, mais il faut se rappeler que le personnel travaille de longues heures, y compris les soirs et les week-ends, et qu'il s'agit de travail sous contrat, et que la rémunération doit tenir compte de ce facteur. Certains avantages sociaux peuvent avoir à être inclus dans le budget. Il ne faut pas oublier que les agents qui estiment qu'ils ne sont pas bien rémunérés tendent à moins s'appliquer, ce qui augmente le risque d'erreurs autres que les erreurs d'échantillonnage. Selon la durée de l'enquête, on peut choisir de payer le personnel à la journée, à la semaine ou au mois. Il est essentiel que les fonds soient disponibles dès le départ, pour payer les traitements et salaires intégralement et en temps voulu. Il faut prévoir des indemnités pour les énumérateurs et les chefs d'équipe qui ont à partir en déplacement. Certains responsables d'enquêtes tendent à payer le personnel de terrain « à l'entrevue effectuée », mais cette pratique peut donner lieu à des abus et est donc déconseillée.

22. Toutes les catégories de personnel, à tous les échelons, doivent être comptées, y compris les personnes qui ne travaillent qu'à temps partiel sur le projet. Le calendrier de l'enquête devrait faciliter l'évaluation du temps de travail de chaque membre de l'équipe.

23. Un tableau de l'emploi du temps du personnel offre également un moyen de déterminer les charges salariales à imputer au budget. Ce tableau utilise également la matrice pour donner un aperçu de l'emploi du temps de chaque membre de l'équipe de l'enquête. Le tableau XIV.3 ci-dessous en donne une illustration. Comme précédemment, il faudra plus de détails dans chaque case pour le processus de planification.

Tableau XIV.3

Matrice du temps de personnel (jours) en regard des activités de l'enquête

Nombre de personnes	Nombre d'hommes-jour pour chaque activité						Total jours
	Consultations	Consultation	Échantillonnage	Travail de terrain	Traitement des données	Rapport	
Directeur							
Superviseur							
Chef d'équipe							
Énumérateur							
Préposé aux données							
Analyste							
Secrétariat							
Chauffeurs							
Autres							
Total (jours)							

24. Les agents de terrain devraient recevoir une indemnité journalière (*per diem*) pour couvrir leurs repas, leurs boissons et autres besoins essentiels pendant leur déplacement. Le montant de cette indemnité devrait se situer dans les limites locales, mais dépasser peut-être le montant habituel, afin de couvrir les cas où la nourriture est insuffisante et pour parer aux urgences.

25. Les frais d'hébergement de tous les agents qui travaillent loin de chez eux doivent être inscrits au budget et payés en temps voulu. Souvent, les agents préfèrent se loger eux-mêmes quand ils vont d'une région à une autre, mais dans d'autres cas, il est plus logique de prendre des dispositions depuis le siège.

2. Transports

26. Les coûts des transports peuvent être estimés avec assez de précision si l'on sait où se trouvent les répondants, c'est-à-dire après que l'échantillon de base a été conçu. Selon les circonstances, on peut conseiller aux énumérateurs d'organiser eux-mêmes leurs voyages, d'enregistrer leurs dépenses pour en demander le remboursement, ou on peut décider d'assurer le déplacement de chaque équipe d'agents de terrain. Cette dernière option est jugée préférable car les membres de l'équipe doivent travailler ensemble, et le chef de l'équipe a ainsi moins de mal à suivre le calendrier d'entrevues. Les coûts additionnels qui ne peuvent être prévus ont trait notamment à la hausse des prix du carburant, à certains phénomènes climatiques rendant les routes impraticables, etc., et ces éventualités sont couvertes par les provisions pour imprévus.

27. Il conviendrait aussi d'inscrire au budget les frais de transport pour les réunions ordinaires des chefs d'équipe avec les responsables de l'enquête.

28. Il peut s'avérer nécessaire d'acheter ou de louer des véhicules/motos/bicyclettes pour le travail de terrain, et il peut être difficile d'inscrire ces dépenses au budget en cas d'inflation.

3. Équipement

29. En général, on peut établir de bonnes estimations des dépenses d'achat d'équipement bien à l'avance. Les problèmes qui peuvent se poser avec ces aspects du budget sont le plus souvent liés à la hausse des prix et à la disponibilité des équipements nécessaires. S'il y a des chances que ces problèmes se posent, il peut être bon d'effectuer ces achats à l'avance et en quantité suffisante pour couvrir tous les besoins de l'enquête. Les matériels d'informatique, de communication, de photocopie et d'impression ne sont pas pris en compte ici.

4. Fournitures

30. Les articles à considérer au titre de ce poste du budget comprennent toutes sortes de fournitures de bureau, logiciels, sacs, cartes, pièces d'identité et panneaux d'affichage, autres articles de bureau, etc. Les fournitures nécessaires à l'impression et à la reproduction représentent une part importante de cette rubrique du budget de l'enquête, car il est essentiel d'avoir accès 24 heures sur 24 à des matériels de copie pendant toute la durée de l'enquête.

5. Autres dépenses

31. On a toujours besoin de publicité et d'informations pendant une enquête. L'ampleur de ces activités dépend entièrement de la nature et de l'envergure de l'enquête, et ces activités peuvent avoir lieu à divers moments tout au long de l'enquête. Il peut s'agir de réunions ou d'ateliers avec tous les intéressés, y compris les chefs communautaires et les usagers, de prises de contact à l'avance avec les répondants, de pu-

blicité, etc. La publicité doit se poursuivre durant toute l'enquête, des informations étant ainsi communiquées aux intéressés en prévision de la diffusion finale des résultats.

32. Pendant certaines phases de l'enquête, il est fait appel à de grands nombres d'agents. Il est essentiel de prévoir suffisamment de place pour la tenue de longues réunions (par exemple, pendant la formation), pour le stockage des questionnaires, pour permettre aux préposés à l'entrée des données et aux superviseurs de travailler confortablement, etc. Parfois, il peut être nécessaire de louer d'autres locaux, par exemple pour se rapprocher du terrain, et parfois, ces locaux-ci peuvent être immédiatement accessibles.

33. La formation peut atteindre des coûts prohibitifs si elle n'a pas été convenablement préparée. Ces coûts comprennent les coûts afférents aux salles de formation et les frais de transport pour les entrevues organisées à titre de formation, plus des indemnités journalières pour tous les intéressés. Tous ces coûts doivent être pris en compte.

34. Il est facile d'oublier de tenir compte de toutes les communications qui sont nécessaires pendant l'enquête : coups de téléphone, courriel, télécopie et poste. Ces dépenses sont souvent difficiles à prévoir au budget, car on ne sait jamais en quelles quantités elles seront nécessaires. En général, on définit une somme forfaitaire, qui est souvent un pourcentage du total, dont on espère qu'elle couvrira toutes ces dépenses. La communication doit être constante avec les équipes de terrain, pour faire face aux imprévus et pour s'assurer du respect des méthodes de collecte des données. Dans les pays où le réseau de téléphone portable est fiable, ces appareils offrent un moyen très utile de communication à distance.

35. Les coûts « occultes » se rapportent à l'infrastructure déjà « en place », telle que les ordinateurs ou les locaux à usage de bureau. Parmi les autres éléments cachés du budget qui ne sont pas aussi évidents figurent les coûts du personnel qui est employé pour l'exécution de tâches se rapportant à plus d'un projet, dont chacun a son propre budget. En règle générale, il est judicieux de tenter d'estimer le temps ou le volume de travail qui sera effectivement consacré à l'enquête, mais on peut aussi parfois estimer grossièrement ces coûts additionnels en pourcentage du coût total. Il est important que tous ces coûts occultes soient identifiés et comptabilisés, afin que, lors de la planification d'enquêtes futures, l'on soit conscient de leur existence et que l'on puisse en tenir compte, même si la situation a changé entre-temps.

6. Exemples d'établissement de budget par postes comptables

36. Comme on l'a indiqué précédemment, il est très difficile d'accéder à l'information sur les coûts effectifs d'une enquête. L'exemple présenté ci-dessous, dû à l'obligeance d'Ajayi (2002), se réfère aux données sur les coûts recueillies auprès de plusieurs pays africains à propos des enquêtes sur les objectifs de la fin de la décennie effectuées en prévision de la demande d'indicateurs de la santé et du bien-être maternels et infantiles présentée par l'Organisation des Nations Unies.

Exemple 2

37. On a obtenu des informations de 12 pays sur les coûts des enquêtes par poste comptable. Les exemples de postes utilisés dans les enquêtes financées sur le budget de l'État sont présentés au tableau XIV.4 ci-dessous, qui indique la proportion du budget total affectée à chaque poste.

Tableau XIV.4.
Coûts par poste comptable en proportion du budget total :
enquêtes sur les objectifs de la fin de la décennie (1999-2000), pour certains pays africains
(en pourcentage)

Pays	Personnel ^a	Transports	Équipement	Fournitures	Autres	Taille de l'échantillon
Angola	62,7	22,2	9,6	1,3	4,2	6 000
Botswana	79,2	0 ^b	10,1	3,5	7,2	7 000
Erytrée	64,0	0 ^b	28,0	4,8	3,2	4 000
Kenya	62,3	22,8	3,3	4,7	6,9	7 000
Lesotho	75,1	5,2	5,8	2,3	11,6	7 500
Madagascar	31,2	6,5	33,3	12,8	16,1	6 500
Malawi	32,0	17,3	23,9	21,6	5,2	6 000
Somalie	43,8	17,7	5,0	1,0	32,5	2 200
Afrique du Sud	69,3	24,0	1,5	3,7	1,5	30 000
Swaziland	29,8	4,3	1,9	1,0	63,0	4 500
République-Unie de Tanzanie	77,9	12,8	1,6	1,2	6,5	3 000
Zambie	81,8	5,2	2,0	5,6	5,4	8 000
Ensemble	62,9	14,9	7,4	6,3	8,5	7 054

Source : Ajayi (2002)

^a Y compris les indemnités journalières.

^b Indique l'impossibilité d'extraire ces informations séparément.

38. Le tableau XIV.4 montre clairement les différences selon les pays dans le budget établi par poste comptable. On pourrait s'attendre à ce qu'un échantillon plus grand se traduise par un plus gros budget de personnel, ce qui ne semble pas être le cas, notamment si l'on compare l'Afrique du Sud à la République-Unie de Tanzanie. Néanmoins, il est probablement vrai que la plupart des enquêtes prévoient d'affecter jusqu'aux deux tiers de leur budget total aux coûts de personnel, y compris aux indemnités journalières pendant le travail de terrain. Pour toute enquête nationale, le deuxième poste par ordre de coût a des chances d'être celui des transports, qui absorbent entre 15 et 20 % du budget total. Le financement de ces enquêtes a été fourni par le Fonds des Nations Unies pour l'enfance (UNICEF) et le gouvernement intéressé, la proportion des coûts financée par l'UNICEF variant d'un pays à un autre.

Exemple 3

39. Cet exemple se rapporte au budget d'une enquête sur les ménages effectuée en 1999 afin d'évaluer l'impact des services aux microentreprises [*Assessing the Impact of Microenterprise Services* (AIMS)] (Barnes et Keogh, 1999; Barnes, 2001). Cette enquête avait pour objet d'étudier les opérations de microfinancement au Zimbabwe et était donc payée par les utilisateurs [financée par Management Systemes International (MSI) à travers l'United States Agency for International Development (USAID)].

40. Le tableau XIV.5 indique une forte proportion du budget (75 %) affectée au personnel, y compris aux indemnités journalières. Cela est dû en partie à la conception de l'enquête, qui était une opération de suivi d'une étude de base effectuée en 1997 et nécessitait la localisation et/ou l'identification des mêmes répondants, ce qui prend beaucoup de temps.

Tableau XIV.5

**Proportion du budget allouée aux différents postes comptables :
évaluation de l'impact des services aux microentreprises [Assessing
the Impact of Microenterprise Services (AIMS)], Zimbabwe (1999) [en pourcentage]**

Personnel	Transports	Fournitures	Autres	Taille de l'échantillon ^a
75	8	912	5	691

^a La taille finale de l'échantillon a été de 599, en raison de l'impossibilité, due à diverses raisons, de localiser 92 des 691 répondants.

D. Principales activités de l'enquête menées dans les limites du cadre budgétaire

41. Une fois que l'on a connaissance de tous les aspects de l'enquête qui nécessiteront un budget, on peut définir les postes comptables à utiliser. Ensuite, on doit considérer les phases de l'enquête et élaborer un budget complet, en utilisant séparément pour chaque phase les postes comptables définis. Cela conduit à l'établissement du cadre budgétaire à l'aide d'une matrice, comme il est indiqué à la section C.

42. Les coûts réels apparaissent à mesure que l'on progresse d'une phase de l'enquête à la suivante, et le budget établi de la même manière rend les comparaisons d'autant plus faciles et permet de rester très sensible aux différences entre le budget et ces coûts.

43. Par ailleurs, cette approche laisse clairement apparaître les liens étroits entre la qualité des données et le calendrier et le budget de l'enquête.

1. Établissement du budget de préparation de l'enquête

44. Cette phase nécessite l'établissement du budget de tous les préparatifs de l'enquête. Il faut examiner tous les postes comptables un à un et estimer exactement ce que chacun nécessitera. Il pourrait être prudent de passer sans tarder commande de fournitures, d'équipements, de réserver des véhicules, etc., si l'on se trouve dans un contexte de forte inflation. Le recrutement de personnel et la publicité sont d'importantes activités, de même que la préparation et la conception de l'échantillon, du questionnaire et des manuels, ainsi que les préparatifs en vue de l'entrée et de la gestion des données.

45. La préparation du cadre d'échantillonnage est un élément important du processus de conception de l'enquête. La nature de ce cadre est fonction du type d'enquête, mais parfois, il faut prévoir beaucoup de temps et/ou de déplacements pour actualiser un cadre existant ou en créer un nouveau. Il faut donc décider des listes à établir, choisir les ménages, les villages ou quelque autre niveau pour l'unité d'échantillonnage, et l'établissement de ces listes exige sa propre allocation budgétaire.

46. La conception des questionnaires, ainsi que la formation et l'établissement des manuels pour le travail de terrain sont d'autres activités qui demandent beaucoup de temps.

2. Établissement du budget d'exécution de l'enquête

47. L'exécution ayant des chances d'être la partie la plus coûteuse de l'enquête, il est très important de calculer avec soin le montant à allouer à chaque poste comptable, en tenant compte de chaque scénario possible. Le temps et le budget alloués à l'impression des questionnaires doivent être étudiés attentivement et planifiés suffi-

samment à l'avance sur la base de sources fiables. En même temps, il ne faut pas perdre de vue que lorsque commence le travail de terrain, le siège central doit se préparer pour l'entrée des données.

48. Comme on l'a souligné précédemment, il ne faut pas rogner sur le travail de terrain pour qu'il tienne dans les limites budgétaires fixées, car on risque, ce faisant, de compromettre la qualité des données et d'accroître le nombre d'erreurs autres que les erreurs d'échantillonnage.

3. Établissement du budget de traitement des données de l'enquête

49. Il faut tenir compte de tous les scénarios possibles lors de l'établissement des budgets d'entrée, de validation, de correction et d'analyse des données, afin d'éviter que ces activités soient bâclées et que les données présentées soient médiocres ou incomplètes. Cette phase s'accompagne d'un gros travail d'impression, et en cherchant à lésiner sur les fournitures, on risque de nuire à la qualité d'ensemble des résultats. Il faut aussi prévoir des moyens informatiques adéquats, y compris des moyens de secours (CD, disquettes).

4. Établissement du budget de préparation du rapport

50. Une fois le travail de terrain terminé et l'entrée des données bien engagée, il faut s'atteler à la phase suivante de l'établissement du budget, à savoir la préparation du rapport et la finalisation. Là encore, la conception de l'enquête joue un rôle, car elle aura permis de déterminer l'ampleur de l'analyse des données et le niveau de détail du rapport. Il est fortement recommandé de tenir une documentation tout au long de l'enquête, car la tenue d'un journal des activités quotidiennes, des décisions, des problèmes et des coûts vient étayer les sections descriptives du rapport. Il faut également examiner de près les postes comptables et allouer des montants adéquats à chacun pour cette phase finale de l'enquête.

5. Exemples d'établissement de budget pour les activités de l'enquête

51. Les informations tirées des exemples de la section C.6 ci-dessus sont présentées ici pour chaque activité.

Exemple 4

52. Si l'on revient à l'exemple 2 (enquêtes sur les objectifs de la fin de la décennie), on dispose d'informations pour déterminer le coût des différentes activités de ces enquêtes dans 10 pays. Le tableau XIV.6 ci-après en donne un résumé.

53. Tous les pays, sauf le Swaziland, indiquent une forte proportion des coûts attribuée à l'exécution de l'enquête : il est probablement raisonnable d'estimer que 70 à 90 % du budget seront consacrés à cette phase de l'enquête. Étant donné que (comme le montre le tableau XIV.4) le Malawi affiche des coûts assez élevés pour les équipements, cela pourrait expliquer la plus forte proportion allouée aux coûts de traitement des données et d'établissement de rapport indiquée au tableau XIV.6. En revanche, on n'a pas d'explication pour la proportion relativement élevée allouée par le Botswana et le Lesotho pour les coûts de traitement des données. C'est pourquoi il a été demandé à ces pays de présenter une « matrice » des coûts indiquant les postes comptables pour

chacune des activités; malheureusement, seules la République-Unie de Tanzanie et l'Érythrée ont présenté une telle récapitulation.

Tableau XIV.6

Coûts des différentes activités d'une enquête en proportion de son budget total : objectifs de la fin de la décennie (1999-2000), pour certains pays africains (en pourcentage)

Pays	Préparation	Exécution ^a	Traitement des données ^b	Rapport ^c	Taille de l'échantillon
Angola	0 ^d	83,6	6,1	10,3	6 000
Botswana	10,4 ^d	59,1	21,7	8,8	7 000
Kenya	0 ^d	93,9	2,6	3,5	7 000
Lesotho	0 ^d	73,2	18,6	8,8	7 500
Madagascar	0,3	78,6	3,0	18,1	6 500
Malawi	5,0	62,7	16,4	15,9	6 000
Afrique du Sud	1,3	93,1	2,9	2,7	30 000
Swaziland	63,0	23,4	7,5	6,1	4 500
République-Unie de Tanzanie	22,7	72,4	3,6	1,3	3 000
Zambie	0,4	92,0	6,4	1,2	8 000
Globalement	7,0	81,0	6,0	6,0	7 054

Source : Ajayi (2002)

^a Y compris la formation, la conception, l'essai pilote et la collecte des données.

^b Y compris l'entrée, la gestion et l'analyse des données.

^c Y compris l'établissement et la diffusion du rapport.

^d Indique l'impossibilité d'extraire cette information séparément.

Exemple 5

54. Pour en revenir à l'exemple 3, nous présentons ci-dessous, au tableau XIV.7, les informations sur les coûts par activité de l'Enquête de 1999 sur l'évaluation de l'impact des services aux microentreprises (AIMS) au Zimbabwe

Tableau XIV.7

Coûts des activités de l'enquête en proportion de son budget total : AIMS Zimbabwe, 1999 (en pourcentage)

Préparation	Exécution ^a	Traitement des données ^b	Rapport ^c	Taille de l'échantillon
4	85	8	3	599

^a Y compris la localisation des répondants, la conception, la formation, la phase pilote et la collecte des données.

^b Y compris l'entrée, la gestion et la correction des données.

^c Se réfère uniquement aux rapports locaux, jusqu'à la production de données corrigées; l'analyse détaillée des données et le rapport final ont fait l'objet de contrats distincts.

55. La proportion sensiblement élevée des coûts d'exécution de l'enquête par rapport au budget total, telle qu'elle ressort du tableau ci-dessus, tient probablement au fait que l'échantillon utilisé pour cette enquête (AIMS) était constitué de 691 répondants qui avaient participé à l'enquête précédente (1997), dont le coût de localisation a été relativement élevé (22 % du budget total).

E. Synthèse

56. Une fois que l'on a établi les coûts à prévoir par chaque poste comptable pour chaque type d'activité, on peut construire une matrice indiquant le budget par activité afin de faciliter l'examen final du budget de l'enquête. Cette matrice aide les responsables de la planification de l'enquête à examiner l'enquête dans son ensemble, à éliminer les discordances et les doubles emplois et à mettre en lumière les principaux coûts à attendre; elle aide également les institutions de financement à comparer les coûts entre diverses enquêtes, ce qui leur permet de mieux évaluer la validité du budget proposé.

57. Comme on l'a indiqué pour l'exemple 4 ci-dessus, deux seulement des 21 pays ayant participé aux enquêtes sur les objectifs de la fin de la décennie ont effectivement produit la matrice demandée, de sorte que, pour cet exemple, on ne peut établir de matrice des postes comptables par activité de l'enquête.

58. Toutefois, on dispose des informations pour l'enquête AIMS et la classification croisée des tableaux XIV.5 et XIV.7 est présentée au tableau XIV.8 ci-dessous.

Tableau XIV.8.

Coûts par poste comptable pour chaque activité de l'enquête en proportion prévue du budget : AIMS Zimbabwe, 1999 (en pourcentage)

	Préparation	Exécution	Données	Rapport	Ensemble
Personnel	3	65	5	2	75
Transports	0	8	0	0	8
Fournitures	0,9	9	2	0,1	12
Autres	0,1	3	1	0,9	5
Ensemble	4,0	85	8	3	100

59. Une matrice telle que celle du tableau XIV.8 ci-dessus, qui indique clairement les besoins budgétaires d'une telle enquête, encourage les institutions de financement à considérer une demande favorablement. De plus, si ces renseignements sont disponibles, on peut plus facilement ajuster le budget pour faire face aux besoins inattendus en période de forte inflation. Enfin, l'enregistrement constant des dépenses, qui doit se poursuivre tout au long de l'enquête, peut facilement s'adapter à une matrice similaire des coûts réels. Bien entendu, il faudra aussi une matrice comme celle qui est présentée ci-dessus, mais avec des montants effectifs en dollars.

60. La donnée finale que l'institution de financement tient à se voir présenter avec un budget proposé est une estimation du coût par ménage ou autre unité d'échantillonnage. Là encore, ce chiffre peut faire fonction d'indication pour un examen réaliste du budget par comparaison avec des exercices similaires.

61. Une telle matrice se prête aisément à des changements dynamiques pendant l'exécution de l'enquête, car elle offre une vue d'ensemble qui permet de voir comment réduire les dépenses dans un domaine pour les accroître dans un autre aux besoins plus pressants, et d'accepter des modifications de la conception, du financement reçu et de l'exécution de l'enquête. Lorsque l'enquête AIMS (1999) a été effectivement exécutée, il a fallu apporter des modifications au budget proposé, principalement dans le domaine des coûts de personnel, en raison d'une forte inflation inattendue. Les responsables de l'exécution de l'enquête ont pu transférer des fonds des postes fournitures, transports et autres catégories liées au travail de terrain, de manière à accroître les crédits alloués aux dépenses de personnel qui étaient justifiées. Le tableau XIV.9 ci-après présente la matrice des dépenses effectives de cette enquête.

Tableau XIV.9

Coûts par poste comptable pour chaque activité de l'enquête en proportion effective du budget : AIMS Zimbabwe, 1999 (en pourcentage)

	Préparation	Exécution	Données	Rapport	Ensemble
Personnel	3,3	69,3	5,6	2,5	80,7
Transports	0	6,6	0	0	6,6
Fournitures	0,6	7,1	2,1	0,1	9,9
Autres	0,1	2,5	0	0,2	2,8
Ensemble	4,0	85,5	7,7	2,8	100

F. Limitations potentielles du budget et pièges

62. Quelle que soit la prudence dont on fait preuve en préparant une enquête, la réalité sur le terrain n'est jamais ce que l'on avait prévu. Il importe d'être conscient à l'avance de cet état de chose, car on peut ainsi inclure des provisions pour imprévus dans la demande de budget finale. Ce poste est généralement défini en pourcentage du coût total, calculé selon les principes recommandés dans les sections précédentes : une provision de 5 à 10 % est généralement jugée acceptable.

63. Indépendamment de l'inclusion d'un pourcentage de provision pour imprévus, il faut être conscient de la situation du pays, surtout si sa situation politique ou économique n'est pas stable. Les institutions de financement doivent être informées des dangers éventuels lorsque le budget leur est soumis, et en restant constamment en contact avec elles durant le cours de l'enquête, on peut rapidement les alerter si des événements menacent le budget de dérapage. Ces dangers peuvent être causés par l'homme ou par l'environnement, et certaines questions liées à la politique locale, à l'économie, au climat, à des mouvements migratoires, etc. peuvent aussi être évoquées dans les contacts permanents avec les bailleurs de fonds et/ou ceux qui ont demandé que soit effectuée l'enquête.

64. Par exemple, lors de l'enquête AIMS de 1999 au Zimbabwe, l'inflation a augmenté régulièrement pendant plusieurs mois, et les coordonnateurs de l'enquête ont pensé qu'ils avaient tenu compte de cette situation lors de l'établissement du budget de l'enquête. Or, au moment même où le travail de terrain devait commencer, les autorités ont gelé le taux de change du dollar des États-Unis à un niveau déraisonnablement bas, ne concordant pas avec le taux d'inflation sans cesse croissant et rendant les coûts prévus totalement irréalistes. Heureusement, Management Systems International s'est montré compréhensif et a autorisé un accroissement des dépenses qui a permis de mener l'exercice à son terme.

65. En pareils cas, il peut être nécessaire de réduire l'effectif et de ne retenir que les éléments les plus efficaces, ou de réduire les coûts par d'autres moyens, par exemple en utilisant des fournitures de bureau plus ordinaires, les transports publics au lieu de véhicules de location, de regrouper les opérations pour réduire les frais généraux, etc. Il est aussi recommandé, si ceux qui financent l'enquête l'autorisent, d'inclure dans la documentation de l'enquête une note déclarant que la portée de l'enquête peut faire l'objet de modifications en cas de circonstances imprévues qui ne permettraient pas, par exemple, de modifier la taille de l'échantillon pour tenir compte de la hausse des prix.

G. Tenue de rapports et résumés

66. On a indiqué précédemment que l'enregistrement journalier des événements durant le cours de l'enquête est indispensable pour suivre toutes les décisions prises et les options envisagées avant de parvenir à ces décisions. Cela comprend l'enregistrement des dépenses.

67. Au stade de la préparation de l'enquête, les coordonnateurs devraient concevoir un ensemble de formulaires à utiliser par tous les employés pour consigner toutes les activités et les dépenses journalières. Ces formulaires devraient indiquer le nombre d'heures de travail, les tâches accomplies, le détail des entrevues, des transports, etc., qui pourraient faire l'objet d'un résumé hebdomadaire. Ainsi, on pourrait à la fois suivre de près le budget et cerner rapidement les problèmes éventuels. En outre, il conviendrait dans toute la mesure du possible d'instituer un système de paiement sur présentation de reçus acceptables.

68. Le suivi et l'établissement de comptes rendus des activités journalières, et de leurs coûts, sont l'une des responsabilités essentielles de la gestion de l'enquête. Il faut différentes formes d'enregistrement pour les différentes phases de l'enquête.

Conception de l'enquête

69. Durant cette phase, le responsable de la gestion de l'enquête doit rester en contact étroit avec toutes les activités, pour en faciliter le suivi. La tenue quotidienne d'un journal est un moyen utile de consigner qui a fait quoi, et cela peut être résumé dans un rapport hebdomadaire. Un enregistrement parallèle des coûts effectifs des transports, des fournitures, de l'hébergement, etc., sera corroboré par le résumé hebdomadaire, de manière à fournir un état des dépenses par semaine. On trouvera des exemples de formulaires d'enregistrement des événements journaliers et hebdomadaires en annexe à ce chapitre.

Exécution de l'enquête

70. Durant l'exécution de l'enquête, le responsable de sa gestion devra pouvoir compter sur l'appui des chefs des équipes de terrain pour lui fournir des comptes rendus de leurs activités journalières ainsi que de leurs dépenses, et les reçus correspondants. Une fois encore, le responsable devra donner un compte rendu hebdomadaire de toutes les dépenses et des journées de travail des membres des équipes, afin que l'on puisse facilement vérifier le pourcentage du budget utilisé. Des exemples des formulaires à utiliser à cette fin sont présentés en annexe.

Établissement de rapports d'enquête

71. Une fois encore, le responsable de la gestion de l'enquête doit rester en contact étroit avec les activités pendant cette phase, et la tenue d'un journal des activités et dépenses quotidiennes permettra d'établir un résumé hebdomadaire. Des formulaires à utiliser à cette fin sont présentés en annexe.

Suivi des dépenses en regard du budget

72. Il est judicieux de confier à une personne la charge du suivi constant des dépenses en regard du budget. Cette personne devrait présenter un compte rendu hebdomadaire des dépenses, comparées aux allocations budgétaires (voir exemple en annexe). Si ce mécanisme est en place dès le début de l'enquête, il sera assez simple de prévoir les problèmes et, s'il y a lieu, de demander des réallocations du budget. Les professionnels des enquêtes doivent être conscients du fait qu'un accroissement du

budget une fois que l'enquête a démarré est un phénomène très inhabituel et que, par conséquent, les ajustements sont la clé du succès de l'enquête.

H. Conclusions

73. Dans ce chapitre, nous avons cherché à donner quelques conseils utiles pour la planification du budget d'une enquête par un examen détaillé de toutes les composantes de l'enquête. Nous avons recommandé une approche dynamique, traitant le budget de deux points de vue, que nous avons illustrée à l'aide d'exemples.

74. Il reste à souligner que cette planification détaillée est d'une importance cruciale si l'on veut mener à bien une étude transparente, fiable et de haute qualité. Tout aussi importante est la nécessité d'un enregistrement quotidien de toutes les activités et actions, qui peut à son tour alimenter le processus comptable et servir de référence sûre pour toute planification d'enquêtes à venir.

Annexe

Exemples de formulaires pour la tenue de listes quotidiennes et hebdomadaires d'activités

Liste individuelle d'activités quotidiennes			
NOM			
Date	Activité	Lieu	Temps passé
Nombre total de jours			

Liste quotidienne d'entrevues				
NOM de l'énumérateur				
Date	Lieu	N° de code de l'entrevue	Temps passé	Résultat de l'entrevue/commentaires

Compte rendu individuel d'activité hebdomadaire			
NOM du chef de l'équipe		Date du rapport	
NOM de l'intéressé	Résumé d'activités	Lieu	Nombre total de jours
Nombre total de jours			

Liste personnelle de dépenses quotidiennes					
NOM					
Date	Lieu	Activité	Détail des dépenses	Montant (dollars)	Reçu n°
Montant total (dollars)					

Liste personnelle de dépenses hebdomadaires					
NOM du chef de l'équipe			Date du rapport		
NOM de l'intéressé	Lieu	Activité	Détail des dépenses	Montant (dollars)	Reçus n ^{os}
Montant total (dollars)					

Relevé hebdomadaire des dépenses						
NOM						
Catégorie	Semaine 1	Semaine 2	Semaine 3	Semaine 4	Semaine 5	Semaine 6
Personnel						
Salaire/ traitement						
Hébergement						
Repas						
Autres						
Transports						
Fournitures						
Autres						

État récapitulatif des dépenses hebdomadaires*								
Catégorie	Budget	Montant cumulé	Semaine 1	Semaine 2	Semaine 3	Semaine 4	Semaine 5	Semaine 6
Personnel								
Salaire/ traitement								
Hébergement								
Repas								
Autres								
Transports								
Carburant								
Véhicules								
Publics								
Autres								
Équipement								
Fournitures								
Autres								
Total à ce jour								

* Peut être tenu sous forme de tableau (par exemple, sur EXCEL).

RÉFÉRENCES

- Ajayi, D. (2002). Communication personnelle.
- Barnes, C. (2001). *An Assessment of Zambuko Trust Zimbabwe*. Assessing the Impact of Microenterprise Services (AIMS) project. Washington, D.C., Management Systems International.
- _____, et E. Keogh (1999). *An Assessment of the Impact of Zambuko's Microfinance Program in Zimbabwe: Baseline Findings*. Assessing the Impact of Microenterprise Services (AIMS) paper. Washington, D.C., Management Systems International.
- Greenfield, T. (1996). *Research Methods: Guidance for Postgraduates*. New York, Arnold, John Wiley and Sons, Inc., p. 306.
- Groves, R. M. (1989). *Survey Errors and Survey Costs*. New York, J. Wiley and Sons, Inc.
- _____, et J. M. Lepkowski (1985). Dual frame mixed mode survey designs. *Journal of Official Statistics*, vol. 1, n° 3, pp. 263-286.
- Nations Unies (1984). *Handbook of Household Surveys* (édition révisée). Studies in Methods, n° 31, numéro de vente : E. 83. XVII.13.
- Scheaffer, R. L., W. Mendenhall et L. Ott (1990). *Elementary Survey Sampling* (4^e éd.). Belmont, Californie, Wadsworth, Inc., p. 97.

Section E

Analyse des données de l'enquête

Introduction

GRAHAM KALTON

Westat

Rockville, Maryland

États-Unis d'Amérique

1. Lorsque les données ont été recueillies, il faut les préparer pour l'analyse. Cette opération se déroule en trois temps. Premièrement, comme on le verra au chapitre XV, il faut décider comment formater les données au mieux pour l'analyse, compte tenu des équipements informatiques disponibles et du logiciel d'analyse utilisé. Souvent, l'analyse porte sur deux unités ou plus : en particulier, dans beaucoup d'enquêtes sur les ménages, ceux-ci et les personnes sont des unités distinctes. Par conséquent, le fichier de données doit pouvoir gérer efficacement des structures hiérarchiques; par exemple, il doit tenir compte du fait que les personnes appartiennent à des ménages et que tous les ménages n'ont pas le même nombre de membres.

2. La deuxième phase de la préparation des données a trait à leur nettoyage ou édition. Immanquablement, les réponses comportent des erreurs identifiables de diverses façons; par exemple, des contradictions par rapport à d'autres réponses ou des impossibilités. Ces erreurs doivent être corrigées avant que puisse commencer l'analyse (voir chapitre XV pour plus de détails sur le nettoyage des données).

3. L'une des tâches importantes du nettoyage des données est de préparer chaque unité sélectionnée pour l'analyse. Toutes les unités sélectionnées doivent être classées dans l'une des trois catégories suivantes : répondant, non-répondant admissible, unité non admissible ou unité au statut d'admissibilité inconnu qui ne répond pas (voir chapitre VIII). Pour être classé comme répondant, il faut généralement plus que la simple présence d'un questionnaire rempli par l'unité. Il faut avoir recueilli un minimum de données acceptables pour que l'unité puisse être classée comme telle. L'attribution d'un statut de répondant nécessite donc plus qu'un examen des questionnaires. Il convient toutefois de noter que, même si une unité est retenue comme répondant pour l'analyse, il peut encore y avoir certains points sur lesquels il n'a pas été obtenu de réponses acceptables. Pour faire face à ce problème, on peut utiliser une certaine forme d'imputation pour affecter des valeurs pour les réponses manquantes.

4. Il faut avoir décidé de l'admissibilité ou non de toutes les unités sélectionnées pour pouvoir passer à la troisième phase de la préparation des données, le calcul des pondérations. Ces pondérations sont calculées pour chacune des unités de l'analyse. Comme avant d'établir des pondérations il faut déterminer les probabilités de sélection de toutes les unités de l'échantillon, il est essentiel de tenir soigneusement des états des probabilités de sélection de ces unités. Les pondérations initiales ou de base des unités sélectionnées se calculent comme l'inverse de leur probabilité de sélection. On ajuste ensuite ces pondérations pour compenser les cas de non-répondants admissibles et pour une certaine proportion de non-répondants au statut d'admissibilité inconnu. Ensuite, on applique souvent un nouvel ajustement pour aligner les distributions pondérées ajustées de certaines variables clés sur les distributions connues de ces

variables obtenues de source extérieure. L'établissement des pondérations est décrit aux chapitres XV et XIX.

5. Lors de la préparation des données, l'une des responsabilités importantes est de veiller à ce que les informations nécessaires pour l'analyse soient consignées dans les données de chaque répondant. Des pondérations sont nécessaires pour chaque unité répondante, pour établir des estimations valables des paramètres de la population étudiée. Il faut des informations sur chaque UPE et chaque strate de chaque unité répondante pour que les erreurs d'échantillonnage puissent être calculées correctement pour les estimations de l'enquête (voir chapitre XXI).

6. Les analyses des données d'une enquête se distinguent à deux titres des analyses décrites dans les manuels de statistiques courants. L'un est la nécessité d'utiliser les pondérations pour compenser l'inégalité des probabilités de sélection, la non-réponse et la non-couverture. En n'utilisant pas ces pondérations, on risque de fausser les estimations des chiffres de la population.

7. L'autre titre est la nécessité de calculer les erreurs d'échantillonnage en tenant compte de la complexité de conception de l'échantillon. Les manuels de statistiques courants supposent en fait un échantillonnage sans restrictions, alors que la plupart des enquêtes sur les ménages utilisent l'échantillonnage stratifié à plusieurs degrés. En général, dans les estimations établies à partir d'un échantillon stratifié à plusieurs degrés, les erreurs d'échantillonnage sont plus grandes que celles établies à partir d'un échantillon rigoureusement aléatoire de même taille, de sorte que l'application des formules des manuels de statistiques courants tend à exagérer la précision des estimations (voir chapitres VI, VII et XXI). Cela veut dire que les logiciels statistiques courants produisent des estimations inexactes des erreurs. Heureusement, toutefois, il existe aujourd'hui un grand nombre de logiciels d'analyse d'enquête qui peuvent établir des estimations appropriées des erreurs d'échantillonnage à partir de données obtenues au moyen d'échantillons complexes. Le chapitre XXI donne un aperçu de plusieurs de ces logiciels.

8. Une bonne part de l'analyse des enquêtes demandées par l'État sont de caractère descriptif. Souvent, les résultats sont présentés sous forme de tableaux, les cases des tableaux indiquant les moyennes, les pourcentages ou les totaux; parfois, ils sont présentés sous forme graphique. En termes étroitement statistiques, les estimations sont souvent très simples, la seule question étant le besoin de s'assurer que des pondérations ont été utilisées. Cependant, il faut tenir compte d'importantes questions de définition et de présentation. Il faut bien veiller à définir le phénomène à mesurer (par exemple, la pauvreté; voir chapitre XVII), et à préciser l'ensemble d'unités sur lesquelles il doit être mesuré. En outre, il faut présenter les résultats d'une façon qui indique clairement ce qui a été mesuré, et pour quel ensemble d'unités. Le chapitre XVI donne des conseils sur la présentation d'estimations descriptives simples.

9. Souvent, le phénomène à mesurer peut être défini d'une façon relativement logique du point de vue des réponses. Parfois, cependant, il est plus complexe et peut nécessiter la création d'un indice utilisant des méthodes statistiques multivariées, telles que l'analyse par grappes et l'analyse de la composante principale. On trouvera plusieurs exemples au chapitre XVIII, y compris, notamment, un exemple pour lequel a été établi un indice de « richesse » à l'aide d'informations sur certaines variables telles que celles de savoir si le ménage avait l'électricité, le nombre de personnes dormant dans une chambre et le principal type d'eau potable.

10. Enfin, il convient de noter que si la production d'estimations descriptives reste la principale forme d'analyse des enquêtes, on utilise de plus en plus des techniques d'analyse des données. On emploie souvent ces techniques pour examiner les

rappports entre les variables et pour rechercher les éventuelles relations de cause à effet. La forme la plus courante de ce type d'analyse est celle où on construit un modèle statistique pour mieux prédire une variable dépendante par rapport à un ensemble de variables indépendantes. Si la variable dépendante est continue (par exemple, le revenu des ménages), on peut utiliser les méthodes de régression linéaire multiple. S'il s'agit d'une variable catégorique à réponse binaire (par exemple, si le ménage a ou n'a pas l'eau courante), on peut alors utiliser les méthodes de régression logistique. Ces méthodes, ainsi que les effets d'un échantillon complexe sur elles, sont décrites aux chapitres XIX et XX. Le chapitre XIX décrit également l'utilisation d'une modélisation à niveaux multiples dans une enquête, et le chapitre XX traite des effets d'un échantillon complexe sur les tests chi-carré des associations entre variables catégoriques.

Chapitre XV

Guide pour la gestion des données d'enquêtes sur les ménages

JUAN MUÑOZ

Sistemas Integrales
Santiago, Chili

RÉSUMÉ

Le présent chapitre décrit le rôle de la gestion des données dans la conception et l'exécution d'enquêtes nationales sur les ménages. Il commence par un examen du rapport entre la gestion des données et la conception du questionnaire, puis il étudie les options passées, présentes et futures concernant l'entrée et l'édition des données et ce qu'elles impliquent pour la gestion générale de l'enquête. Les sections suivantes contiennent des directives pour la définition des critères de contrôle de qualité et l'élaboration de programmes d'entrée des données d'enquêtes nationales complexes sur les ménages, jusqu'à la diffusion des ensembles de données de l'enquête. La dernière section est consacrée à un examen du rôle de la gestion des données en tant qu'appui à la conception de l'échantillon.

Termes clés : contrôle de cohérence; nettoyage des données, édition des données, gestion des données, enquête sur les ménages, critères de contrôle de qualité.

A. Introduction

1. Si l'on a souvent souligné l'importance de la gestion des données dans les enquêtes sur les ménages, cette gestion est généralement perçue comme un ensemble de tâches liées à la phase de tabulation des résultats, c'est-à-dire comme des activités menées vers la fin de l'enquête, à l'aide d'ordinateurs dans les bureaux du siège, sous l'autorité d'analystes et de programmeurs.

2. Cette vision restrictive est en train de changer. L'expérience des deux dernières décennies a montré que la gestion des données peut et doit jouer un rôle critique dès les premiers stades de l'enquête. Par ailleurs, il est clair à présent que cette gestion ne s'achève pas avec la publication des premiers rapports statistiques.

3. La manifestation la plus claire de l'importance de la gestion des données avant même les phases d'analyse a été donnée par l'Étude de la mesure des niveaux de vie effectuée par la Banque mondiale et par d'autres enquêtes qui ont réussi à intégrer le contrôle informatique de qualité aux opérations sur le terrain. Même lorsque l'entrée des données ne fait pas partie du travail de terrain, il importe que les responsables de la gestion des données participent à la conception des questionnaires, afin de s'assurer que les unités statistiques observées sont convenablement reconnues et identifiées, que les instructions données aux enquêteurs sur les questions à sauter sont

explicites et correctes et que des redondances sont délibérément incorporées aux questionnaires pour les contrôles de cohérence.

4. À la fin du calendrier d'exécution de l'enquête, l'idée que le produit fini est une publication imprimée contenant un ensemble de tableaux statistiques a laissé la place au concept d'une base de données qui peut non seulement être utilisée par l'office de statistiques pour l'établissement de tableaux, mais aussi être accessible aux chercheurs, aux décideurs et au public en général. Le rapport sommaire décrivant les résultats de l'enquête n'est plus perçu comme l'étape finale, mais plutôt comme le point de départ de tout un ensemble de travaux analytiques qui peuvent se prolonger de nombreuses années après la clôture du projet et la dissolution de l'équipe de l'enquête.

5. Le présent chapitre commence par un examen des rapports entre la gestion des données de l'enquête et la conception du questionnaire, suivi d'une étude des options passées, présentes et futures concernant l'entrée et l'édition des données et ce qu'elles impliquent pour la gestion générale de l'enquête. Les sections suivantes contiennent des directives pour la définition des critères de contrôle de qualité et l'élaboration de programmes d'entrée des données d'enquêtes nationales complexes sur les ménages, jusqu'à la diffusion des ensembles de données de l'enquête comprise. La dernière section est consacrée à un examen du rôle de la gestion des données en tant qu'appui à la conception de l'échantillon

B. Gestion des données et conception du questionnaire

6. La gestion des données de l'enquête commence dès la conception du questionnaire et peut grandement influencer celle-ci. Le responsable de la gestion des données doit être consulté sur chaque projet de questionnaire, car sa perspicacité permet de discerner les défauts dans la définition des unités, les questions à sauter, etc. Dans la présente section, nous examinerons certains des aspects du questionnaire qui nécessitent une attention particulière à ce stade.

7. *Nature et identification des unités statistiques observées.* Chaque enquête sur les ménages a pour objet de recueillir des informations sur une unité statistique importante – le ménage – ainsi que sur une variété d'unités secondaires au sein du ménage – les personnes, les postes budgétaires, les parcelles de terrain, les cultures, etc. Le questionnaire doit être clair et explicite quant à la nature exacte de ces unités, et il doit donner à chaque unité observée son propre moyen d'identification.

8. L'identification du ménage lui-même apparaît généralement sur la page de couverture du questionnaire. Parfois, elle est constituée d'une longue série de chiffres et de lettres qui représentent l'emplacement géographique et les procédures d'échantillonnage utilisées pour choisir le ménage. Si elle peut paraître évidente, l'utilisation de tous ces codes d'identification des ménages doit être vérifiée avec le plus grand soin, car elle est laborieuse, sujette à erreurs et coûteuse (il faut souvent 20 chiffres ou plus pour identifier les quelques centaines de ménages de l'échantillon); parfois, elle ne garantit même pas que l'identification est exclusive à une seule unité, par exemple lorsque les codes géographiques figurant sur la page de couverture identifient l'unité d'habitation mais ne tiennent pas compte de la possibilité que plusieurs ménages y habitent. Il est plus facile et plus sûr d'identifier les ménages au moyen d'un simple numéro de série qui peut être inscrit à la main ou tamponné sur la page de couverture du questionnaire, voire imprimé à l'avance. L'emplacement géographique, la condition urbaine/rurale, les codes d'échantillonnage et le reste des données de la page de couverture deviennent des attributs importants du ménage qui, à ce titre, doivent figurer dans les ensembles de données de l'enquête, mais pas nécessairement aux fins d'identification. Un bon compromis entre ces deux extrêmes (la liste de tous les codes

d'échantillonnage détaillés et un simple numéro de série donné au ménage) serait de donner un numéro de série de trois ou quatre chiffres aux unités primaires d'échantillonnage (UPE) utilisées pour l'enquête, puis un numéro de série de deux chiffres aux ménages de chaque UPE.

9. La nature des unités statistiques secondaires est souvent évidente (par exemple, les individus composant le ménage), mais il peut se présenter des ambiguïtés lorsque ce qui semble être une unité est en fait un ensemble d'unités de types différents. Tel peut être le cas, par exemple, lorsqu'un homme à qui l'on a demandé de décrire la principale activité de son emploi a plusieurs activités tout aussi importantes les unes que les autres ou exerce plus d'un emploi pendant une période de référence donnée. Ou lorsqu'une femme à qui on a demandé le sexe ou le poids de son dernier enfant a eu des jumeaux de poids différents. Toutefois, bien que de telles situations doivent pouvoir être évitées par une bonne conception et un essai pilote du questionnaire, elles se produisent souvent de façon subtile, et c'est là que la perspicacité du responsable de la gestion des données intervient pour aider les spécialistes à les détecter.

10. Quelle que soit leur nature, les unités secondaires composant le ménage doivent toujours être spécifiquement identifiées. Cela peut se faire au moyen de codes numériques attribués par l'enquêteur, mais il est généralement préférable, dans la mesure du possible, que ces codes soient imprimés à l'avance sur le questionnaire.

11. *Redondances délibérées.* Lors de la conception du questionnaire, on peut envisager d'y inclure délibérément des redondances afin de détecter les fautes de l'enquêteur ou les erreurs d'entrée des données. Les exemples les plus courants de ces redondances sont les suivants :

- Ajouter une ligne pour les « totaux » au bas des colonnes qui indiquent des montants. L'établissement de ces totaux incombe souvent à l'enquêteur, mais même lorsque tel n'est pas le cas, leur inclusion est utile car elle offre un moyen très efficace (souvent le seul) de détecter les erreurs d'entrée des données ou les omissions. En fait, les totaux peuvent être ajoutés au bas des colonnes de chiffres aux fins de contrôle de la qualité, même si la somme de ces chiffres ne représente pas une mesure significative de l'ampleur (par exemple, on peut ajouter un total au bas d'une colonne indiquant des quantités, pas des montants, de produits alimentaires achetés, même si cela veut dire que l'on additionne des chiffres hétérogènes, par exemple des kilos de pain à des kilos de pommes de terre, ou même à des litres de lait). Ce point est approfondi lors de l'examen des vérifications typographiques ci-dessous.
- Ajouter un caractère de contrôle aux codes de certaines variables importantes (telles que la profession ou l'activité d'une personne ou la nature d'un poste de consommation). Ce caractère est un chiffre ou une lettre qui peut être déduit du reste des chiffres du code par des opérations arithmétiques effectuées à l'entrée des données. Voici un algorithme courant du caractère de contrôle : multiplier le dernier chiffre du code par 2, le précédent par trois, etc. (si le code a plus de six chiffres, répéter la suite de multiplicateurs 2, 3, 4, 5, 6, 7), et ajouter les résultats. Le caractère de contrôle est la différence entre cette somme et le plus proche multiple supérieur de 11 (le chiffre 10 est représenté par la lettre K). Les algorithmes du caractère de contrôle sont construits de telle manière que les erreurs de codage les plus courantes, telles que la transposition ou l'omission de caractères, produit un caractère de contrôle erroné.

C. Stratégies opérationnelles pour l'entrée et l'édition des données

12. Nombreux sont les cas où les activités d'entrée et d'édition des données recueillies dans le cadre d'enquêtes sur les ménages s'effectuent après coup à l'échelon central; toutefois, d'autres enquêtes intègrent déjà l'entrée des données aux opérations sur le terrain. Dans un proche avenir, on pourrait envisager la conduite d'entrevues assistées par ordinateur. La présente section analyse les répercussions des diverses stratégies sur l'organisation, de même que les traits communs et les spécificités des logiciels d'entrée et d'édition des données conçus pour chacune de ces options.

13. *Entrée des données centralisée.* L'entrée des données à l'échelon central était la seule option connue avant l'apparition des micro-ordinateurs, et reste encore la norme pour de nombreuses enquêtes. Selon cette option, l'entrée des données est considérée comme un processus industriel qui doit se dérouler dans le cadre d'ateliers centralisés, une fois les entrevues terminées. Il s'agit de convertir la matière brute (les informations figurant sur les questionnaires) en un produit intermédiaire (fichiers lisibles sur machine) qui ont besoin de nouvelles améliorations (apportées au moyen de programmes d'édition et d'opérations de secrétariat) pour que l'on dispose comme produit fini d'une base de données « propre ».

14. Durant la phase initiale d'entrée des données, les priorités sont la rapidité et l'obtention de fichiers contenant des informations reflétant parfaitement les informations recueillies sur les questionnaires. En fait, les préposés à l'entrée des données ne sont pas censés « réfléchir » à ce qu'ils font et doivent se contenter de recopier fidèlement les données qui leur sont communiquées. Parfois, les questionnaires sont soumis à une entrée de données « aveugle » pour garantir que tel est bien le cas.

15. Jusqu'au milieu des années 1970, les données étaient entrées sur des machines spécialisées aux capacités très limitées. Aujourd'hui, cette opération est presque toujours effectuée à l'aide de micro-ordinateurs qui peuvent être programmés pour permettre des contrôles de qualité, cette capacité étant rarement utilisée. L'impression générale est qu'il faut limiter les contrôles de qualité lors de l'entrée des données car les opérateurs ne sont pas qualifiés pour apporter des corrections si une erreur est constatée. De plus, la détection d'erreurs et leur correction ralentit le processus d'entrée des données. D'après ce courant de pensée, le contrôle de qualité doit être laissé exclusivement pour le processus d'édition.

16. *Entrée des données sur le terrain.* À partir du milieu des années 1980, l'intégration du contrôle informatique de qualité aux opérations de terrain est apparue comme l'un des moyens essentiels d'améliorer la qualité et l'actualité des enquêtes sur les ménages. Ces idées ont été développées d'abord dans le cadre des Enquêtes de la Banque mondiale pour la mesure des niveaux de vie (LSMS), et ont été appliquées par la suite à diverses enquêtes complexes sur les ménages. Ainsi, l'entrée des données et les contrôles de cohérence sont appliqués à chaque ménage dans le cadre des opérations de terrain, et les erreurs et incohérences sont résolues au moyen de nouvelles visites aux ménages.

17. L'intégration a pour avantage principal et le plus direct d'améliorer la qualité de l'information, car elle permet de corriger les erreurs et incohérences pendant que les enquêteurs sont encore sur le terrain plutôt que par un « nettoyage » ultérieur au siège. Ce nettoyage est long et laborieux, et produit au mieux des bases de données qui sont intérieurement homogènes mais ne reflètent pas nécessairement les réalités observées sur le terrain. Le doute découle de la quantité de décisions – généralement non étayées – à prendre à distance, et longtemps après la collecte des données.

18. L'intégration de contrôles informatiques de qualité permet aussi de produire en temps voulu, généralement à peine quelques semaines après la fin des opérations sur le terrain, des bases de données prêtes pour la tabulation et l'analyse. En fait, les bases de données peuvent être préparées alors même que l'enquête se déroule, ce qui permet aux responsables de suivre effectivement les opérations sur le terrain.

19. L'intégration offre également l'avantage de favoriser l'application de critères uniformes par tous les enquêteurs et pendant toute la période de collecte des données, ce qui était difficile à réaliser auparavant. En fait, l'ordinateur devient un assistant incorruptible et infatigable des superviseurs de l'enquête.

20. L'intégration des contrôles informatiques de qualité aux opérations sur le terrain a également diverses répercussions sur l'organisation de l'enquête, dont la plus importante est qu'elle nécessite que le personnel de terrain soit organisé en équipes. L'équipe de terrain est généralement dirigée par un superviseur et comprend un préposé à l'entrée des données en plus de deux à quatre enquêteurs.

21. L'organisation des opérations sur le terrain est fonction des moyens technologiques disponibles. Les deux options les plus courantes sont l'ordinateur de bureau et l'ordinateur bloc-notes et comprennent les opérations suivantes :

- Le préposé à l'entrée des données travaille sur un ordinateur de bureau dans un endroit précis (généralement le bureau régional de l'office de statistique) et organise le travail de terrain de manière que le reste de l'équipe visite chaque endroit désigné (en général, une unité primaire d'échantillonnage) au moins deux fois, afin de lui laisser le temps d'entrer les données et d'en vérifier la cohérence entre les deux visites. Lors de la deuxième visite et des visites ultérieures, les enquêteurs reposent les questions pour lesquelles des erreurs, omissions ou incohérences ont été détectées par le programme d'entrée des données;
- Le préposé à l'entrée des données travaille sur un ordinateur bloc-notes et accompagne le reste de l'équipe lors des visites. L'équipe tout entière reste sur place jusqu'à ce que les données soient entrées et certifiées complètes et correctes par le programme d'entrée des données.

22. Ces deux options ont des exigences extérieures qui doivent être dûment prises en compte par les responsables de la planification et de la gestion de l'enquête. L'une d'entre elles suppose l'accès à l'électricité pour les ordinateurs, ce qui peut être problématique dans les pays où l'électrification est peu développée. Si l'on utilise des ordinateurs de bureau dans des lieux fixes, on peut avoir à installer des groupes électrogènes et à veiller à qu'ils disposent constamment de mazout. Si l'on utilise à la place des blocs-notes, on peut avoir à se munir de panneaux solaires portables.

23. La différence évidente entre ces deux options est que, pour que le contrôle informatique de qualité soit intégré au travail de terrain, il faut que l'entrée et l'édition des données soient établies et déboguées avant que démarre l'enquête. Lorsque l'entrée des données est centralisée, cela est aussi utile (afin de permettre que l'entrée des données se déroule parallèlement aux opérations de terrain) mais pas absolument nécessaire.

24. *Entrevues menées sans l'usage de questionnaires sur papier.* L'utilisation d'ordinateurs portables pour éliminer tout simplement les questionnaires sur papier est très séduisante en raison des avantages de l'automatisation de certaines parties de l'entrevue, telles que les instructions données pour sauter certaines questions. Toutefois, bien que cette technologie existe depuis près de 20 ans, on a fait très peu pour appliquer sérieusement cette méthode aux enquêtes complexes sur les ménages dans les pays en développement. En fait, même dans les offices nationaux de statistique les plus

avancés, la pratique de l'interrogation sans questionnaire sur papier n'a été limitée jusqu'ici qu'à des exercices relativement simples, telles que les enquêtes sur l'emploi et la collecte de prix pour le calcul de l'indice des prix à la consommation.

25. La principale raison de cet état de chose est que bien que l'interrogation sans questionnaire sur papier se prête à des entrevues qui suivent un cours linéaire, avec un début et une fin, de nombreuses enquêtes sur les ménages effectuées dans les pays en développement et les pays en transition nécessitent parfois de multiples visites à chaque ménage, des entrevues séparées avec chaque membre de la famille, et d'autres procédures qui sont moins strictement structurées.

26. Malgré l'absence d'une réelle expérience concrète, on peut faire certaines observations sur ce dont il faut tenir compte lors de la conception et de l'exécution d'une entrevue sans questionnaire sur papier :

- Dans certains cas, l'interface avec le programme d'entrée des données consiste en une série de questions apparaissant l'une après l'autre sur l'écran de l'ordinateur, mais dans d'autres cas, il est nécessaire de reproduire visuellement les questionnaires en montrant en même temps plusieurs champs d'entrée des données. Il semble que cela soit particulièrement important pour les modules sur les dépenses et la consommation, où l'enquêteur a besoin de « voir » simultanément plusieurs postes de consommation. L'interface doit également permettre de marquer les questions en cas de doute, et il doit prévoir la possibilité de retourner auprès du ménage pour une seconde entrevue sans avoir à répéter toutes les questions.
- En général, la conception du questionnaire est un processus qui nécessite plusieurs mois de travail et fait intervenir différentes personnes (des spécialistes de la question traitée, des praticiens de l'enquête, etc.). Avec un questionnaire sur papier, ce processus comprend la préparation, la distribution, l'analyse et l'expérimentation de diverses versions successives du questionnaire jusqu'à l'accord sur la version finale. Il reste encore à définir les opérations équivalentes pour quelque chose qui ne figurera jamais sur papier.
- La formation des enquêteurs devra être revue en fonction de la nouvelle technologie. Nous savons comment former des enquêteurs à l'utilisation de questionnaires sur papier (formation théorique, simulations, entrevues à titre d'essai, manuels didactiques, etc.), mais peu de travaux ont été consacrés à concevoir des techniques équivalentes pour les enquêtes sans questionnaire sur papier.
- Enfin, des méthodes de supervision efficaces ont été mises au point. Un vaste ensemble de procédures (vérification visuelle des questionnaires, entrevues de vérification, etc.) ont vu le jour au cours des 50 dernières années pour vérifier le travail accompli sur le terrain par les enquêteurs. Toutes ces procédures étaient fondées sur la notion de questionnaire sur papier et doivent à présent être repensées pour des entrevues sans de tels questionnaires. Il est très probable que ces nouvelles technologies offriront des possibilités totalement différentes, et beaucoup plus performantes, de supervision efficace; en effet, la plupart des ordinateurs portables offrent des possibilités d'enregistrer la voix qui pourraient servir pour enregistrer automatiquement certaines parties de l'entrevue choisies au hasard en même temps que les fichiers de données. L'addition d'un système mondial de positionnement pourrait également permettre d'enregistrer automatiquement le lieu et l'heure de l'entrevue.

D. Critères de contrôle de qualité

27. Quelle que soit la stratégie retenue pour le contrôle de qualité, les données inscrites sur le questionnaire doivent être soumises à cinq types de vérifications : vérifications de l'ordre de grandeur, vérifications par rapport à des données de référence, vérifications des questions à sauter, vérifications de la cohérence et vérifications typographiques. Ici, nous examinerons la nature de ces contrôles et la façon dont ils peuvent s'effectuer dans les divers cadres opérationnels possibles.

28. Les vérifications de l'ordre de grandeur ne visent qu'à s'assurer que les valeurs de chaque variable ne s'inscrivent que dans certaines limites de validité. Les variables catégoriques ne peuvent avoir que l'une des valeurs prédéfinies dans le questionnaire (par exemple, le code identifiant le sexe ne peut être que « 1 » pour les hommes et « 2 » pour les femmes); les variables chronologiques doivent indiquer des dates valables, et les variables numériques doivent se situer entre le minimum et le maximum prescrit (par exemple, 0 à 95 ans pour l'âge).

29. On se trouve en présence d'un cas particulier de vérifications de l'ordre de grandeur lorsque l'on peut vérifier les données de deux domaines étroitement liés par rapport à des tableaux de référence. Voici certaines situations courantes à cet égard :

- *Cohérence des données anthropomorphiques.* Ici, les grandeurs enregistrées pour la taille, le poids et l'âge sont vérifiées par comparaison avec les tables de référence de l'Organisation mondiale de la santé. Toute valeur donnée pour les indicateurs habituels (taille à un certain âge, poids à un certain âge, poids pour une certaine taille) qui s'écarte de plus de trois écarts types de la norme doit être signalée comme une erreur possible, afin d'être mesurée de nouveau.
- *Cohérence des données relatives à la consommation d'aliments.* Ici, les valeurs enregistrées pour le code alimentaire, la quantité achetée et le montant payé sont vérifiées par rapport à un tableau spécifique des prix unitaires possibles.

30. Même lorsque les données sont entrées dans des lieux centralisés, il est généralement utile de détecter et de corriger les erreurs d'ordre de grandeur dès la phase initiale de l'entrée des données, plutôt que d'attendre la phase de l'édition, car les erreurs d'ordre de grandeur sont souvent le résultat de l'opération d'entrée des données elle-même plutôt que d'erreurs de la part de l'enquêteur. Un signe indiquant l'erreur, tel qu'un bip ou un champ clignotant sur l'écran, peut être déclenché lorsqu'une donnée s'écarte de l'ordre de grandeur normal. S'il ne s'agit que d'une faute de frappe, le préposé à l'entrée des données peut la corriger immédiatement. Toutefois, il devrait être possible de passer outre à ce signal si la grandeur entrée correspond à ce qui figure sur le questionnaire. Dans ce cas, l'erreur doit être signalée, afin que le préposé puisse la corriger par la suite en vérifiant le questionnaire (ou que l'enquêteur repose la question lors d'une seconde entrevue, si les données sont entrées sur le terrain). Entre-temps, la donnée suspecte peut être conservée dans un format spécial signalant qu'elle est douteuse.

31. *Vérifications des questions à sauter.* Il s'agit ici de vérifier s'il n'a pas été commis d'erreurs concernant les questions à sauter. Par exemple, une simple vérification permet de s'assurer que les questions à ne poser qu'aux écoliers n'ont pas de réponse de la part d'un enfant qui a répondu non à une question initiale sur la scolarisation. Il est plus compliqué de s'assurer que chaque répondant a rempli les modules appropriés du questionnaire. Selon son âge et son sexe, chaque membre de la famille est censé répondre à certaines sections du questionnaire et en sauter d'autres. Par exemple, les enfants de moins de cinq ans doivent être mesurés dans la section des questions anthro-

pomorphiques, mais il est évident qu'il ne faut pas leur poser les questions concernant la profession. Les questions sur la fécondité ne sont à poser qu'aux femmes âgées de 15 à 49 ans, mais pas aux hommes.

32. À l'avenir, les entrevues assistées par ordinateur (sans questionnaires sur papier) pourraient devenir chose courante pour les enquêtes effectuées dans les pays en développement, et le programme lui-même devrait permettre de contrôler le bon choix des questions à sauter, du moins dans certains cas. Toutefois, dans les autres cadres opérationnels (entrée des données dans un lieu central ou sur le terrain), le programme d'entrée des données ne devrait pas suivre de lui-même la configuration des questions à sauter. Par exemple, s'il est répondu non à la question : « Êtes-vous inscrit dans une école ? », les champs dans lesquels entrer les données sur le type d'école fréquentée, la classe, etc., doivent toujours être soumis au préposé à l'entrée des données. Si des réponses sont effectivement portées sur le questionnaire, elles peuvent alors être entrées, mais le programme signalera une erreur concernant des questions à sauter. Le superviseur ou l'enquêteur (ou le personnel au bureau central) peut déterminer ultérieurement la nature de l'erreur. Il est très possible que le non enregistré ait été une erreur. Si le programme d'entrée des données avait automatiquement sauté les champs suivants, l'erreur n'aurait pas été corrigée ou redressée.

33. *Vérifications de la cohérence.* Ces vérifications permettent de s'assurer que les chiffres donnés en réponse à une question sont conformes à ceux donnés en réponse à une autre. Il se produit une vérification simple lorsque les deux chiffres émanent de la même unité, par exemple la date de naissance et l'âge d'une personne donnée. Ces vérifications sont plus compliquées lorsqu'il s'agit de comparer des données de deux unités d'observation différentes ou plus.

34. Il n'y a pas de limite naturelle au nombre de vérifications de cohérence. Les versions bien écrites du programme d'entrée des données pour une enquête complexe sur les ménages peuvent en compter plusieurs centaines. Toutefois, étant donné que le temps imparti pour concevoir les programmes d'entrée et d'édition des données est toujours limité (il est en général d'environ deux mois), il faut de la compétence et du bon sens pour décider exactement des vérifications à prévoir. Certaines vérifications de cohérence applicables à presque toutes les enquêtes sur les ménages se sont révélées particulièrement efficaces et sont devenues pratiquement la norme. Ces vérifications portent sur :

- *La compatibilité démographique du ménage.* On vérifie la compatibilité entre les âges de tous les membres de la famille en fonction de leurs relations de parenté. Par exemple, les parents doivent être d'au moins 15 ans plus âgés que leurs enfants, les époux doivent être de sexe différent, etc;
- *La compatibilité des professions.* La présence ou l'absence de certaines sections doit être compatible avec les professions déclarées par chacun des membres de la famille. Par exemple, la section travaux agricoles ne doit figurer que si certains membres de la famille ont indiqué « agriculteurs » dans la section « travail »;
- *Compatibilité entre l'âge et les autres caractéristiques.* On peut vérifier que l'âge de chaque personne est compatible avec ses caractéristiques personnelles, telles que sa situation de famille, sa relation avec le chef de famille, la classe dans laquelle il est (pour les enfants d'âge scolaire) ou la dernière classe suivie (pour ceux qui ont quitté l'école). Par exemple, un enfant de huit ans ne doit pas être dans une classe plus élevée que la troisième année d'école primaire;

- *Dépenses.* Ici, plusieurs vérifications de compatibilité sont possibles. Il ne doit y avoir de chiffres positifs concernant les dépenses de consommation de la famille pour des articles tels que manuels scolaires et frais de scolarité que si plusieurs membres de la famille déclarent qu'un enfant est scolarisé. De même, seules les familles qui ont l'électricité doivent indiquer des dépenses d'électricité.
- *Totaux de contrôle.* Comme on l'a vu précédemment, le fait d'ajouter un total de contrôle chaque fois que des chiffres peuvent s'additionner est un moyen pratique de s'assurer de la bonne conception du questionnaire.

35. *Vérifications typographiques.* Dans les débuts du traitement informatique des enquêtes, la vérification des fautes de frappe était pratiquement le seul contrôle de qualité qui était effectué à l'entrée des données. Pour ce faire, on faisait en sorte que les données de chaque questionnaire soient entrées deux fois par des personnes différentes. Ces procédures « aveugles » sont rarement utilisées aujourd'hui, car les autres vérifications de compatibilité aujourd'hui possibles les rendent inutiles. Cela dit, dans certains cas, elles peuvent encore servir.

36. L'erreur typographique type est celle de l'inversion de deux chiffres (par exemple, entrer « 14 » au lieu de « 41 »). Une telle erreur sur l'âge peut être décelée par des vérifications de compatibilité avec la situation de famille ou les relations familiales. Par exemple, le questionnaire concernant un adulte marié ou veuf de 41 ans dont l'âge a été entré par erreur comme 14 ans apparaît avec un signal d'erreur lors de la vérification de l'âge puis de celle de la situation de famille. En revanche, la même erreur dans les dépenses mensuelles de viande peut facilement échapper à l'attention, car \$14 ou \$41 pourraient être tous deux des montants plausibles.

37. Cela montre combien il importe de tenir compte des perspectives de gestion des données dans la phase de conception du questionnaire. Par exemple, les totaux de contrôle peuvent réduire sensiblement les erreurs typographiques, car demander à l'enquêteur de faire la somme des chiffres déclarés à l'aide d'une calculatrice de poche équivaut aux procédures « aveugles » de vérification des données. Des chiffres de contrôle peuvent également être utilisés à cette fin pour certaines variables importantes. On peut également appliquer ces procédures pour entrer les données de certaines parties du questionnaire, mais il est à la fois inutile et peu pratique de le faire pour l'ensemble du questionnaire; entre autres raisons parce que les méthodes modernes d'entrée des données sont généralement fondées sur l'intervention d'un seul préposé et non de deux.

E. Conception du programme d'entrée des données

38. La conception d'un bon programme d'entrée et d'édition des données d'une enquête est à la fois une technique et un art. Dans la présente section, nous examinerons certaines des plates-formes qui permettent aujourd'hui de faire face aux aspects techniques du processus de conception et certaines des questions plus subtiles liées à la conception d'interfaces pour les préposés à l'entrée des données et les futurs utilisateurs des ensembles de données d'enquêtes.

39. *Plates-formes de conception.* Il existe beaucoup de plates-formes de conception de programmes d'entrée et d'édition de données sur le marché, mais il y en a peu qui soient spécifiquement adaptées aux exigences de gestion de données des enquêtes complexes sur les ménages. Il ressort d'une étude effectuée au milieu des années 1990 par la Banque mondiale qu'à l'époque deux plates-formes DOS convenaient : le programme d'Études pour la mesure du niveau de vie (LSMS) conçu par la Banque mondiale et le programme de Système intégré de traitement sur microprocesseur (IMPS)

d'United States Bureau of the Census. Ces deux plates-formes ont progressé depuis cette étude, s'adaptant à l'évolution du matériel et des systèmes d'exploitation. IMPS a été remplacé par le « Census and Survey Processing System (CSPPro), application fondée sur Windows et qui offre quelques possibilités de tabulation, en plus de son rôle principal de programme d'entrée et d'édition de données. Le programme LSMS a laissé la place à LSD-2000, fondé sur Excel, qui sert à la fois à la conception du questionnaire et du programme d'entrée des données.

40. CSPPro et LSD-2000 (et leurs prédécesseurs) ont montré qu'ils pouvaient aider à concevoir les programmes d'entrée et d'édition de données d'enquêtes nationales complexes sur les ménages de beaucoup de pays. Ces plates-formes sont en outre faciles à acquérir et à utiliser. N'importe quel programmeur, en vérité quiconque a des rudiments d'informatique, devrait pouvoir acquérir en l'espace de deux semaines l'aptitude technique nécessaire pour entreprendre de concevoir un programme opérationnel d'entrée de données

41. *Principes de conception.* Malheureusement, ces plates-formes de conception ne peuvent conseiller les programmeurs sur le point de savoir exactement quel type de programme d'entrée de données concevoir. On peut même faire valoir que la facilité d'utilisation de ces plates-formes risque de rendre trop aisée la conception de programmes inadéquats. Il importe que les responsables de la gestion de l'enquête évitent de confondre la maîtrise des outils avec leur bon choix d'utilisation en associant à la fois des programmes expérimentés et des spécialistes du sujet traité à la conception des programmes d'entrée et d'édition des données. À cet égard, certaines directives pratiques peuvent se révéler utiles :

- *Conception de l'écran d'entrée des données.* Ces écrans devraient ressembler le plus possible aux pages correspondantes du questionnaire, mais il y a de nombreuses exceptions à cette règle. Par exemple, si le questionnaire présente les questions personnelles sous la forme d'une matrice (avec les questions comme rangées et les membres de la famille comme colonnes, ou inversement), il est généralement préférable de préparer un écran d'entrée de données distinct pour chaque personne, plutôt que de reproduire la grille du questionnaire sur l'écran. L'une des raisons en est que le nombre de répondants varie; une autre, que les unités statistiques observées sont les personnes et non pas les familles.
- *Faire la distinction entre les situations impossibles et les situations improbables.* Bien entendu, le programme d'entrée des données doit signaler comme erreur toute situation qui représente une impossibilité logique ou naturelle (par exemple, une fille plus âgée que sa mère), mais il doit aussi réagir aux situations qui ne sont pas naturellement impossibles mais très improbables (par exemple, une fille plus jeune de moins de 15 ans que sa mère). En principe, le programme d'entrée des données devrait évaluer la gravité des erreurs et réagir différemment selon le degré de gravité, au même titre qu'un humain qui vérifierait visuellement le questionnaire. Malheureusement, certains programmeurs ne consacrent pas suffisamment d'efforts à cette question. Un signe édifiant à cet égard est la tendance à toujours définir le maximum possible d'une variable quantitative comme « 999 » (autant de 9 que le champ d'entrée des données le permet). L'inefficacité de cette pratique est évidente : certes, les champs d'entrée des données doivent être suffisamment vastes pour pouvoir recevoir les chiffres les plus élevés possibles, mais les maxima doivent être suffisamment raisonnables pour que toute valeur improbable soit signalée comme telle.

- *Langage utilisé pour signaler les erreurs.* Certains des critères de qualité inclus dans le programme d'entrée des données permettent de signaler les erreurs détectées par des moyens relativement simples qui n'ont pas besoin d'explications ou sont faciles à comprendre. Par exemple, le programme LSMS signale les valeurs qui sortent du champ des possibilités en affichant des flèches clignotantes pointées vers le haut « ↑↑↑ » ou vers le bas « ↓↓↓ » selon que la valeur en question est jugée trop faible ou trop élevée. Cependant, les contrôles plus complexes nécessitent une indication plus claire ou plus explicite. Par exemple, une vérification de cohérence des données sur le ménage pourrait produire un avertissement du type : « Attention, Lucy (Code ID 05, fillette de 9 ans) a peu de chances d'être la fille de Mary (Code ID 02, femme de 21 ans) », de préférence imprimé sur papier plutôt que n'apparaissant que sur l'écran de l'ordinateur. Ce type de programmation intelligente peut prendre plus de temps que les options apparemment plus simples (telles que l'utilisation de codes d'erreur), mais il permet d'économiser de nombreuses heures de travail sur le terrain et de formation du personnel de terrain, et il dégage les programmeurs de l'obligation d'établir un code d'erreurs.
- *Codes d'identification des variables.* En général, une enquête complexe sur les ménages utilise des centaines de variables. Les programmeurs chargés du programme d'entrée des données ont besoin de désigner ces variables par des codes, selon les conventions de la plate-forme de conception utilisée. Pour ce faire, il importe de choisir un système de codage rationnel et simple dès le début du processus de conception du programme d'entrée des données, car cela facilite la communication entre les membres de l'équipe de conception et cela permet de gagner du temps durant les phases successives de préparation et de diffusion des ensembles de données de l'enquête. Cela dit, il peut être plus difficile qu'il ne semble de trouver un bon système de codage. Le processus, simple au début, les premiers codes à établir désignant des variables telles que « l'AGE », « le SEXE » et ainsi de suite, peut vite devenir très compliqué à mesure qu'il devient de plus en plus difficile de trouver des codes mnémoniques. Il est souvent simple de se référer aux numéros de section et de question du questionnaire, sans que les codes soient nécessairement faciles à identifier (par exemple, si « l'âge » et « le sexe » sont les variables 4 et 5 de la section 1, leur code pourrait être « S1Q4 » et « S1Q5 », respectivement).
- *Entrée des données.* Lorsque l'entrée des données est intégrée aux opérations sur le terrain, l'unité la plus naturelle pour l'entrée des données est le ménage. En effet, le préposé à l'entrée des données n'a jamais plus qu'un ou quelques questionnaires sur lesquels travailler, et les contrôles de cohérence et la notification d'erreurs s'effectuent ménage par ménage. Lorsque l'entrée des données se fait à un échelon central, le préposé peut se voir confier des blocs de 10 à 20 ménages (correspondant par exemple à des localités ou à des UPE). L'idée est que : a) l'entrée des données d'un bloc doit être confiée à un seul préposé travaillant sur un même ordinateur pendant deux jours au plus; et que b) l'ensemble de questionnaires correspondants doit être facile à garder en mémoire et à extraire à tout moment.

F. Organisation et diffusion des ensembles de données de l'enquête

42. La structure des ensembles de données d'une enquête doit refléter la nature des unités statistiques observées. Autrement dit, les données d'une enquête complexe sur les ménages **ne peuvent pas** être mises en mémoire sous la forme présentée au ta-

bleau XV.1 ci-après, c'est-à-dire d'un simple fichier rectangulaire avec une rangée pour chaque ménage et une colonne pour chacun des champs du questionnaire.

Tableau XV.1.

Données d'une enquête sur les ménages mises en mémoire dans un simple fichier rectangulaire

	Variable 1	Variable 2	...	Variable j	...	Variable m	
Ménage 1				
Ménage 2				
...	
Ménage <i>i</i>			...	Donnée <i>i, j</i>	...		
...	
Ménage <i>n</i>				

43. Une telle structure (appelée « fichier plat ») conviendrait si toutes les questions se référaient au ménage en tant qu'unité statistique, mais comme on l'a vu précédemment, tel n'est pas le cas. Certaines des questions se réfèrent à des unités statistiques secondaires qui figurent en nombres variables dans chaque ménage, comme les personnes, les cultures, les articles de consommation et ainsi de suite. L'indication de l'âge et du sexe de chaque membre de la famille comme variables au niveau du ménage serait à la fois un gaspillage (car le nombre de variables nécessaires serait fonction de la taille de la famille la plus nombreuse plutôt que la taille moyenne des familles) et extrêmement compliquée pour l'analyse (car même les tâches simples telles que l'obtention de la répartition par âge et par sexe obligerait à sonder un nombre variable de combinaisons âge-sexe au sein de chaque famille).

44. Les plates-formes CPro et LSD-2000 utilisent toutes deux une structure de fichier qui s'accommode bien des complexités découlant de l'étude de nombreuses unités statistiques différentes, tout en réduisant les besoins de mise en mémoire et en s'adaptant bien au logiciel statistique au stade de l'analyse.

45. La structure des données maintient une correspondance rigoureuse entre chaque unité statistique observée et les informations figurant dans les fichiers informatiques, utilisant un type d'enregistrement différent pour chaque type d'unité statistique. Par exemple, pour gérer les données figurant sur une liste de ménages, un type de registre serait défini pour les variables de la liste, et les données correspondant à chaque individu seraient entrées dans un registre particulier de ce type. De même, pour le module de consommation alimentaire, un type de registre correspondrait aux aliments et les données correspondant à chaque aliment particulier seraient entrées dans des registres particuliers de ce type.

46. Le nombre de registres de chaque type peut varier. Cela économise l'espace dans la mémoire, car les registres n'ont pas besoin de toujours prévoir un espace correspondant au cas nécessitant le registre le plus grand.

47. En principe, un seul type de registre est nécessaire pour chaque unité statistique, mais parfois, pour des raisons pratiques, on peut avoir à définir plus d'un type de registre pour la même unité. Par exemple, les questions sur l'éducation et la santé peuvent être enregistrées sur deux registres séparés, même si l'unité statistique est la même personne dans les deux cas.

48. Chaque registre est identifié par un code en trois parties ou plus. La première partie désigne le « type de registre », et apparaît au début de chaque registre. Elle indique si les informations proviennent, par exemple, de la page de couverture, du module santé ou des dépenses alimentaires. Le type de registre est suivi, dans tous les cas, par le numéro du ménage. Dans la plupart des types de registre, un troisième code d'identification est nécessaire pour faire la distinction entre des unités statistiques du même type au sein du ménage, par exemple le numéro d'identification de la personne ou le code correspondant à la dépense. Dans quelques cas, il n'y a qu'une unité pour le niveau d'observation, et le troisième code d'identification n'est pas nécessaire. Par exemple, les caractéristiques du logement ne sont généralement indiquées que pour un logement par ménage. Dans quelques cas, il peut y avoir un quatrième code. Par exemple, le troisième peut servir à identifier l'entreprise du ménage et le quatrième, chaque équipement détenu par l'entreprise.

49. Les codes d'identification sont suivis des données proprement dites enregistrées dans le cadre de l'enquête pour chaque unité particulière, enregistrées dans des champs de longueur fixe du même ordre que pour les questions du questionnaire. Toutes les données sont enregistrées sous le code normalisé ASCII (*American Standard Code for Information Interchange*).

50. Les ensembles de données de l'enquête doivent être organisés uniquement sous la forme de fichiers plats séparés (un pour chaque type d'information) en vue de leur diffusion, car le format de champ de longueur fixe de la structure d'origine convient également pour le transfert des données sur des Systèmes de gestion de base de données (DBMS) aux fins de traitement complémentaire ou sur un logiciel de statistique normalisé aux fins de tabulation et d'analyse. Le transfert des données sur DBMS est très facile car la structure d'origine se traduit presque directement en un format de base de données normalisé (DBF) qui est accepté par tous comme éléments d'entrée pour les tableaux (ici, les codes d'identification servent de liens naturels entre les tableaux).

51. La diffusion nécessite également que la structure de chaque type d'information soit convenablement documentée dans un Code de l'enquête qui doit être donné à tout utilisateur désireux de travailler avec les ensembles de données. Ce code doit indiquer clairement la position et la longueur de chaque variable enregistrée. Pour les variables catégoriques, il doit également spécifier le codage. La figure XV.1 ci-après présente une page du Code de l'Enquête sur les niveaux de vie au Népal (les codages de certaines variables ont été abrégés).

52. Les plates-formes CPro et LSD-2000 permettent de produire le code comme sous-produit du processus de conception du programme d'entrée de données. LSD-2000 offre également une interface pour convertir les fichiers d'entrée des données en fichiers DBF et pour transférer les données dans les logiciels les plus communément utilisés (Ariel, CPro, SAS, SPSS et STATA). Cela montre combien il importe de définir avec soin le système de codage des variables durant la phase de conception du programme d'entrée des données : si tel est le cas, les analystes peuvent immédiatement utiliser les données de l'enquête dès que les ensembles de données deviennent disponibles.

Figure XV.1
Enquête sur les niveaux de vie au Népal II

Record Type 002		Section 1, Part A1 : Liste des ménages				
VARIABLE	CODE	RT	DE	LONGUEUR	TYPE	
Ménage		2	4	5	QNT	
CODE ID	IDC	2	9	2	QNT	
1 Nom	Q01	2	11	24	TYP	
D Ethnicité	Q01A	2	35	3		Chhatri 001
						Brahmin Hill 002
					
						Autres 102
2 Sexe	Q02	2	38	1		Hommes 1
						Femmes 2
3. Relation	Q03	2	39	2		Chef 1
						Conjoint 2
						Enfant 3
					
						Autre parent 11
						Serviteur/ de la famille du serviteur 12
						Locataire/de la famille du locataire 13
						Autre personne n'appartenant pas à la famille 14
4A District de naissance	Q04A	2	41	2	QLN	Tapejung 01
						Panchthar 02
					
						Autre pays 93
4B District de naissance U/R	Q04B	2	43	1	QLN	Urbain 1
						Rural 2
5. Âge	Q05	2	44	2	QNI	
6 Situation de famille	Q06	2	46	1	QLN	Marié 1
						Divorcé 2
						Séparé 3
						Veuf 4
						Célibataire 5
7 Conjoint sur la liste ?	Q07	2	47	1	QLN	Oui 1
						Non 2
8 Code ID du conjoint	Q08	2	48	2	QNT	
9 Mois passés au domicile	Q09	2	50	2	QNT	
10 Membre ou non ?	Q10	2	52	1	QLN	Oui 1
						Non 2

G. Gestion des données durant le processus d'échantillonnage

53. La présente section traite du rôle de la gestion des données dans la conception et la mise en œuvre d'échantillons pour enquêtes sur les ménages. Elle contient des recommandations pour l'informatisation des cadres d'échantillonnage et pour les

premiers stades de la sélection de l'échantillon, y compris certaines méthodes pratiques de stratification implicite et de sélection d'unités primaires d'échantillonnage (UPE) avec probabilité proportionnelle à la taille (PPS). Elle analyse l'élaboration d'une base de données avec l'avant-dernière unité d'échantillonnage comme sous-produit des stades d'échantillonnage précédents, et met l'accent sur le rôle de cette base de données comme outil de gestion lors de l'exécution de l'enquête sur le terrain, et sur la façon dont son contenu peut être actualisé à l'aide des informations recueillies sur le terrain (notamment des résultats de l'établissement de la liste de ménages et des données concernant la non-réponse) afin de produire les pondérations à utiliser au stade de l'analyse.

54. *Organisation du cadre d'échantillonnage de la première phase.* Pour de nombreuses enquêtes sur les ménages, les unités d'échantillonnage de la première phase sont les zones d'énumération (ZE) du recensement définies lors du dernier recensement national. La création d'un fichier informatique de la liste de toutes les ZE du pays est un moyen pratique et efficace d'établir le cadre d'échantillonnage de la première phase. À l'exception des pays où le nombre de ZE est élevé (comme le Bangladesh, qui en compte plus de 80 000), le meilleur moyen de procéder est d'utiliser un logiciel de tableur comme Excel, avec une rangée pour chaque ZE, et des colonnes pour toutes les informations nécessaires. Ce tableur doit donner l'identification complète de la ZE et une indication de sa taille (par exemple, sa population, ou le nombre de ménages ou de logements qu'elle regroupe). Il est généralement plus pratique de créer un tableur pour chaque strate de l'échantillon. La figure XV.2 ci-après montre à quoi un cadre d'échantillonnage de première phase pourrait ressembler parmi les strates d'un pays hypothétique (l'écran Excel a été divisé en deux fenêtres pour montrer en même temps la première et la dernière ZE).

55. Dans cet exemple, les 1 326 ZE de la strate Forêt sont identifiées par des codes géographiques et les noms des divisions administratives du pays (provinces ou districts) et par un numéro de série à l'intérieur de chaque district. Le cadre d'échantillonnage indique également le nombre de ménages et la population de chaque ZE au moment du recensement, et précise si la ZE est urbaine ou rurale.

56. Avant de passer aux étapes suivantes de la sélection de l'échantillon, il est essentiel de s'assurer que le cadre d'échantillonnage est complet et correct en vérifiant les chiffres de population avec les totaux du recensement publiés par l'office de statistiques. Il importe aussi de vérifier que la taille de toutes les ZE est suffisamment grande pour qu'elles puissent être utilisées comme unités primaires d'échantillonnage. Si l'échantillon exige des grappes d'avant-dernière phase de 25 ménages chacune, par exemple, cette condition ne pourra être remplie par les ZE de moins de 25 ménages. Dans ce cas, il convient de combiner les petites ZE avec des ZE voisines pour former les unités primaires d'échantillonnage. Ce processus peut s'avérer laborieux s'il faut rechercher les ZE voisines à la main, en se référant continuellement aux cartes du recensement. Toutefois, comme les offices de statistiques attribuent souvent des numéros de série aux ZE selon un certain critère géographique (l'ordonnement en spirale), de telle sorte que les ZE qui sont voisines sur le tableur sont également voisines sur le terrain, il est généralement possible d'effectuer automatiquement les groupements sur le tableur. Dans notre exemple, chaque ZE compte plus de 30 ménages, et aucun groupement n'est nécessaire. Il convient toutefois de noter que l'exemple ci-dessus permettrait difficilement de suivre cette procédure car les ZE urbaines et rurales sont mélangées dans la liste numérique, situation qui a peu de chances de se produire dans la réalité. Autrement dit, le groupement de ZE voisines par ordinateur ne peut se faire lorsque les ZE urbaines et rurales sont mélangées sur la liste plutôt que groupées par catégorie.

Figure XV.2
Utilisation d'un tableau comme cadre d'échantillonnage de la première phase

The image shows a Microsoft Excel spreadsheet with two worksheets. The first worksheet, 'Atlantis Sample Frame.xls:1', contains a table with columns for Province, Province Name, Ward, Ward Name, CEA, Urban/Rural, Households, and Population. The data rows represent different wards in West Tazenda, Macondo. The second worksheet, 'Atlantis Sample Frame.xls:2', contains a similar table for Barzakul, Povai. The status bar at the bottom of each worksheet indicates the current selection: 'Forest' in the first and 'West Hills' in the second.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Province	Province Name	Ward	Ward Name	CEA	Urban/Rural	Households	Population						
2	1	West Tazenda	207	Macondo	01	R	41	256						
3	1	West Tazenda	207	Macondo	02	U	59	328						
4	1	West Tazenda	207	Macondo	03	R	57	288						
5	1	West Tazenda	207	Macondo	04	U	50	320						
6	1	West Tazenda	207	Macondo	05	R	58	300						
7	1	West Tazenda	207	Macondo	06	U	52	276						
8	1	West Tazenda	207	Macondo	07	U	37	238						
9	1	West Tazenda	207	Macondo	08	R	32	180						
10	1	West Tazenda	207	Macondo	09	R	46	250						

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1318	8	Barzakul	414	Povai	15	R	73	254						
1319	8	Barzakul	414	Povai	16	R	87	308						
1320	8	Barzakul	414	Povai	17	U	44	205						
1321	8	Barzakul	414	Povai	18	R	77	386						
1322	8	Barzakul	414	Povai	19	R	79	342						
1323	8	Barzakul	414	Povai	20	U	79	331						
1324	8	Barzakul	414	Povai	21	U	35	154						
1325	8	Barzakul	414	Povai	22	R	42	203						
1326	8	Barzakul	414	Povai	23	U	101	474						
1327	8	Barzakul	414	Povai	24	R	59	313						
1328														
1329														

57. Avant la première phase de l'échantillonnage, il faut aussi décider si le cadre d'échantillonnage a besoin d'être conçu selon certains critères afin de stratifier implicitement l'échantillon à l'intérieur de chaque strate explicite. Les divisions administratives sont presque toujours utilisées à cette fin, mais dans certains cas, on peut estimer qu'un autre critère, par exemple une stratification urbaine/rurale, est encore plus important. Si l'on suppose que, dans notre exemple, tel est le cas, le cadre d'échantillonnage doit être divisé entre urbain et rural, puis par province, puis par district et enfin selon le numéro de série de la ZE. Cela peut se faire aisément avec la commande « tri » prévue par le programme de tableur de la figure XV.3.

Figure XV.3

Exécution d'une stratification implicite

The screenshot shows two worksheets in Microsoft Excel. The first worksheet, 'Atlantis Sample Frame.xls:1', contains data for West Tazenda and Dabuli wards. The second worksheet, 'Atlantis Sample Frame.xls:2', contains data for Barzakul ward. The data is organized into columns for Province, Province Name, Ward, Ward Name, CEA, Urban/Rural, Households, and Population.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Province	Province Name	Ward	Ward Name	CEA	Urban/Rural	Households	Population						
2	1	West Tazenda	207	Macondo	02	U	59	328						
3	1	West Tazenda	207	Macondo	04	U	50	320						
4	1	West Tazenda	207	Macondo	06	U	52	276						
5	1	West Tazenda	207	Macondo	07	U	37	238						
6	1	West Tazenda	207	Macondo	11	U	68	357						
7	1	West Tazenda	207	Macondo	12	U	40	236						
8	1	West Tazenda	207	Macondo	17	U	53	312						
9	1	West Tazenda	207	Macondo	19	U	70	331						
10	1	West Tazenda	211	Dabuli	02	U	50	290						

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1318	8	Barzakul	414	Povai	06	R	169	531						
1319	8	Barzakul	414	Povai	09	R	107	382						
1320	8	Barzakul	414	Povai	12	R	136	416						
1321	8	Barzakul	414	Povai	14	R	115	411						
1322	8	Barzakul	414	Povai	15	R	73	254						
1323	8	Barzakul	414	Povai	16	R	87	308						
1324	8	Barzakul	414	Povai	18	R	77	386						
1325	8	Barzakul	414	Povai	19	R	79	342						
1326	8	Barzakul	414	Povai	22	R	42	203						
1327	8	Barzakul	414	Povai	24	R	59	313						
1328														
1329														

58. *Sélection des unités primaires d'échantillonnage selon une probabilité proportionnelle à la taille.* La plupart des enquêtes sur les ménages choisissent les unités primaires d'échantillonnage en utilisant une probabilité proportionnelle à la taille (PPS). Lorsqu'il figure dans le cadre d'échantillonnage, le nombre de ménages de la ZE est généralement utilisé comme mesure de sa taille, mais dans certains cas, on peut utiliser la population ou le nombre de logements. Nous illustrerons à présent la procédure PPS et supposerons que la conception exige la sélection de 88 ZE avec une probabilité proportionnelle au nombre de ménages (colonne G du tableur) dans la strate Forêt (voir figure XV.4).

59. Premièrement, créer une nouvelle colonne sur le tableur, indiquant la taille cumulée des ZE. Entrer la formule $=I1+G2$ dans la cellule I2 et la copier jusque dans la dernière rangée de la colonne I (on remarquera que la dernière rangée dans la colonne I indique le nombre total de ménages de la strate Forêt (110,388)).

Figure XV.4
Sélection d'un échantillon PPS (première étape)

The image shows a Microsoft Excel spreadsheet with two worksheets. The first worksheet, 'Atlantis Sample Frame.xls:1', contains the following data:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Province	Province Name	Ward	Ward Name	CEA	Urban/Rural	Households	Population						
2	1	West Tazenda	207	Macondo	02	U	59	328	59					
3	1	West Tazenda	207	Macondo	04	U	50	320	109					
4	1	West Tazenda	207	Macondo	06	U	52	276	161					
5	1	West Tazenda	207	Macondo	07	U	37	238	198					
6	1	West Tazenda	207	Macondo	11	U	68	357	266					
7	1	West Tazenda	207	Macondo	12	U	40	236	306					
8	1	West Tazenda	207	Macondo	17	U	53	312	359					
9	1	West Tazenda	207	Macondo	19	U	70	331	429					
10	1	West Tazenda	211	Dabuli	02	U	50	290	479					

The second worksheet, 'Atlantis Sample Frame.xls:2', contains the following data:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1318	8	Barzakul	414	Povai	06	R	169	531	109,613					
1319	8	Barzakul	414	Povai	09	R	107	382	109,720					
1320	8	Barzakul	414	Povai	12	R	136	416	109,856					
1321	8	Barzakul	414	Povai	14	R	115	411	109,971					
1322	8	Barzakul	414	Povai	15	R	73	254	110,044					
1323	8	Barzakul	414	Povai	16	R	87	308	110,131					
1324	8	Barzakul	414	Povai	18	R	77	386	110,208					
1325	8	Barzakul	414	Povai	19	R	79	342	110,287					
1326	8	Barzakul	414	Povai	22	R	42	203	110,329					
1327	8	Barzakul	414	Povai	24	R	59	313	110,388					
1328														
1329														

60. Deuxièmement, créer une autre colonne avec la taille cumulée mise à l'échelle des ZE en multipliant les valeurs de la colonne I par le coefficient de mise à l'échelle de 88/110 388 (l'idée est d'avoir une colonne qui augmente de zéro au nombre de ZE à sélectionner, proportionnellement à la taille des ZE; voir figure XV.5). Entrer la formule $=I2*88/110388$ dans la cellule J2 et la copier jusque dans la dernière rangée de la colonne J.

Figure XV.5
Sélection d'un échantillon PPS (deuxième étape)

The screenshot shows two Excel worksheets. The top worksheet, 'Atlantis Sample Frame.xls:1', has the following data:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Province	Province Name	Ward	Ward Name	CEA	Urban/Rural	Households	Population						
2	1	West Tazenda	207	Macondo	02	U	59	328	59	0.05				
3	1	West Tazenda	207	Macondo	04	U	50	320	109	0.09				
4	1	West Tazenda	207	Macondo	06	U	52	276	161	0.13				
5	1	West Tazenda	207	Macondo	07	U	37	238	198	0.16				
6	1	West Tazenda	207	Macondo	11	U	68	357	266	0.21				
7	1	West Tazenda	207	Macondo	12	U	40	236	306	0.24				
8	1	West Tazenda	207	Macondo	17	U	53	312	359	0.29				
9	1	West Tazenda	207	Macondo	19	U	70	331	429	0.34				
10	1	West Tazenda	211	Dabuli	02	U	50	290	479	0.38				

The bottom worksheet, 'Atlantis Sample Frame.xls:2', has the following data:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1318	8	Barzakul	414	Povai	06	R	169	531	109,613	87.38				
1319	8	Barzakul	414	Povai	09	R	107	382	109,720	87.47				
1320	8	Barzakul	414	Povai	12	R	136	416	109,856	87.58				
1321	8	Barzakul	414	Povai	14	R	115	411	109,971	87.67				
1322	8	Barzakul	414	Povai	15	R	73	254	110,044	87.73				
1323	8	Barzakul	414	Povai	16	R	87	308	110,131	87.80				
1324	8	Barzakul	414	Povai	18	R	77	386	110,208	87.86				
1325	8	Barzakul	414	Povai	19	R	79	342	110,287	87.92				
1326	8	Barzakul	414	Povai	22	R	42	203	110,329	87.95				
1327	8	Barzakul	414	Povai	24	R	59	313	110,388	88.00				
1328														
1329														

61. Troisièmement, entrer un chiffre choisi au hasard entre 0 et 1 dans la première cellule à partir du haut d'une nouvelle colonne et ajouter ce chiffre à toutes les cases de la colonne J afin de créer une nouvelle colonne aux chiffres ainsi décalés (voir figure XV.6). On peut choisir des chiffres au hasard automatiquement à partir du tableur, mais il est préférable de prendre ces chiffres à l'extérieur (par exemple, en utilisant un tableau de chiffres aléatoires) pour empêcher le système de choisir un échantillon différent après chaque calcul. Par exemple, entrer le chiffre choisi au hasard de 0,73 de la case K1, puis entrer la formule =J2 + K\$1 dans la case K2 et la copier dans toute la colonne K.

Figure XV.6
Sélection d'un échantillon PPS (troisième étape)

The image shows two Excel worksheets. The first worksheet, 'Atlantis Sample Frame.xls:1', contains the following data:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Province	Province Name	Ward	Ward Name	CEA	Urban/Rural	Households	Population			0.73			
2	1	West Tazenda	207	Macondo	02	U	59	328	59	0.05	0.78			
3	1	West Tazenda	207	Macondo	04	U	50	320	109	0.09	0.82			
4	1	West Tazenda	207	Macondo	06	U	52	276	161	0.13	0.86			
5	1	West Tazenda	207	Macondo	07	U	37	238	198	0.16	0.89			
6	1	West Tazenda	207	Macondo	11	U	68	357	266	0.21	0.94			
7	1	West Tazenda	207	Macondo	12	U	40	236	306	0.24	0.97			
8	1	West Tazenda	207	Macondo	17	U	53	312	359	0.29	1.02			
9	1	West Tazenda	207	Macondo	19	U	70	331	429	0.34	1.07			
10	1	West Tazenda	211	Dabuli	02	U	50	290	479	0.38	1.11			

The second worksheet, 'Atlantis Sample Frame.xls:2', contains the following data:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1318	8	Barzakul	414	Povai	06	R	169	531	109,613	87.38	88.11			
1319	8	Barzakul	414	Povai	09	R	107	382	109,720	87.47	88.20			
1320	8	Barzakul	414	Povai	12	R	136	416	109,856	87.58	88.31			
1321	8	Barzakul	414	Povai	14	R	115	411	109,971	87.67	88.40			
1322	8	Barzakul	414	Povai	15	R	73	254	110,044	87.73	88.46			
1323	8	Barzakul	414	Povai	16	R	87	308	110,131	87.80	88.53			
1324	8	Barzakul	414	Povai	18	R	77	386	110,208	87.86	88.59			
1325	8	Barzakul	414	Povai	19	R	79	342	110,287	87.92	88.65			
1326	8	Barzakul	414	Povai	22	R	42	203	110,329	87.95	88.68			
1327	8	Barzakul	414	Povai	24	R	59	313	110,388	88.00	88.73			
1328														
1329														

62. L'échantillon est défini par les rangées où la partie entière du chiffre décalé change. Dans cet exemple, le changement se fait au passage de 0,97 à 1,02 pour la ZE 17 du district 207 (Macondo) de la province numéro 1 (West Tazenda), ce qui implique que cette ZE est la première à être choisie pour l'échantillon. Le nouveau changement se fait au passage de 1,99 à 2,09 pour la ZE 01 du district 226 (Balayan) de la même province, qui est donc la deuxième à être choisie. On peut signaler automatiquement l'échantillon sélectionné en entrant la formule $=\text{INT}(K2)-\text{INT}(K1)$ dans la case L2 et en la copiant dans toute la colonne L. L'échantillon est défini par les cases de la colonne L n'ayant pas un zéro (voir figure XV.7).

63. La liste de toutes les unités sélectionnées lors de la première étape doit être reportée sur une feuille de calcul séparée qui deviendra un outil essentiel pour la gestion de l'enquête. Par exemple, les responsables de cette gestion peuvent ajouter des colonnes afin d'inscrire des détails concernant toutes les principales activités menées dans chaque UPE (les dates prévues et effectives des travaux de terrain et d'entrée des données, l'identification de l'équipe responsable, etc.).

Figure XV.7.
Sélection d'un échantillon PPSA (quatrième étape)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Province	Province Name	Ward	Ward Name	CEA	Urban/Rural	Households	Population			0.73			
2	1	West Tazenda	207	Macondo	02	U	59	328	59	0.05	0.78	0		
3	1	West Tazenda	207	Macondo	04	U	50	320	109	0.09	0.82	0		
4	1	West Tazenda	207	Macondo	06	U	52	276	161	0.13	0.86	0		
5	1	West Tazenda	207	Macondo	07	U	37	238	198	0.16	0.89	0		
6	1	West Tazenda	207	Macondo	11	U	68	357	266	0.21	0.94	0		
7	1	West Tazenda	207	Macondo	12	U	40	236	306	0.24	0.97	0		
8	1	West Tazenda	207	Macondo	17	U	53	312	359	0.29	1.02	1		
9	1	West Tazenda	207	Macondo	19	U	70	331	429	0.34	1.07	0		
10	1	West Tazenda	211	Dabuli	02	U	50	290	479	0.38	1.11	0		
11	1	West Tazenda	211	Dabuli	03	U	81	370	560	0.45	1.18	0		
12	1	West Tazenda	211	Dabuli	06	U	77	258	637	0.51	1.24	0		
13	1	West Tazenda	211	Dabuli	07	U	60	252	697	0.56	1.29	0		
14	1	West Tazenda	211	Dabuli	08	U	43	312	740	0.59	1.32	0		
15	1	West Tazenda	211	Dabuli	10	U	75	303	815	0.65	1.38	0		
16	1	West Tazenda	211	Dabuli	11	U	62	311	877	0.70	1.43	0		
17	1	West Tazenda	211	Dabuli	12	U	52	291	929	0.74	1.47	0		
18	1	West Tazenda	211	Dabuli	13	U	63	297	992	0.79	1.52	0		
19	1	West Tazenda	211	Dabuli	15	U	62	345	1,054	0.84	1.57	0		
20	1	West Tazenda	211	Dabuli	17	U	53	289	1,107	0.88	1.61	0		
21	1	West Tazenda	211	Dabuli	18	U	60	308	1,167	0.93	1.66	0		
22	1	West Tazenda	211	Dabuli	19	U	57	325	1,224	0.98	1.71	0		
23	1	West Tazenda	211	Dabuli	20	U	59	311	1,283	1.02	1.75	0		
24	1	West Tazenda	211	Dabuli	21	U	57	319	1,340	1.07	1.80	0		
25	1	West Tazenda	211	Dabuli	24	U	59	347	1,399	1.12	1.85	0		
26	1	West Tazenda	211	Dabuli	27	U	51	331	1,450	1.16	1.89	0		
27	1	West Tazenda	211	Dabuli	29	U	58	328	1,508	1.20	1.93	0		
28	1	West Tazenda	211	Dabuli	30	U	78	384	1,586	1.26	1.99	0		
29	1	West Tazenda	226	Balayan	01	U	125	483	1,711	1.36	2.09	1		
30	1	West Tazenda	226	Balayan	05	U	41	247	1,752	1.40	2.13	0		
31	1	West Tazenda	226	Balayan	06	U	75	369	1,827	1.46	2.19	0		

64. La feuille de calcul sera utilisée notamment pour calculer les probabilités de sélection et les coefficients correspondants (ou pondérations) nécessaires pour tirer des estimations non faussées de l'échantillon. La feuille de calcul simplifiée présentée ci-dessous ne nécessite pas de séparation par strate. Il vaut mieux inscrire toutes les UPE sélectionnées sur une même feuille en indiquant la strate dans l'une des colonnes. La figure XV.8 présente la feuille indiquant l'échantillon des 19 premières des 88 ZE sélectionnées de notre exemple.

Figure XV.8
Tableau indiquant les unités primaires d'échantillonnage sélectionnées

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
	Province	Province Name	Ward	Ward Name	CEA	Urban/Rural	Households	Population	Stratum					
2	1	West Tazenda	207	Macondo	17	U	53	312	Forest					
3	1	West Tazenda	226	Balayan	01	U	125	483	Forest					
4	1	West Tazenda	226	Balayan	53	U	69	394	Forest					
5	1	West Tazenda	226	Balayan	90	U	43	192	Forest					
6	1	West Tazenda	242	Haliyal	52	U	48	279	Forest					
7	1	West Tazenda	255	Gronau	15	U	52	333	Forest					
8	1	West Tazenda	259	Pazar	04	U	79	395	Forest					
9	1	West Tazenda	401	Tolbo	21	U	84	361	Forest					
10	1	West Tazenda	401	Tolbo	38	U	130	463	Forest					
11	2	East Tazenda	267	Xanadu	06	U	125	511	Forest					
12	2	East Tazenda	267	Xanadu	25	U	105	424	Forest					
13	2	East Tazenda	270	Quetta	02	U	166	580	Forest					
14	2	East Tazenda	270	Quetta	21	U	138	407	Forest					
15	2	East Tazenda	270	Quetta	45	U	177	485	Forest					
16	2	East Tazenda	275	Mosken	10	U	150	368	Forest					
17	2	East Tazenda	275	Mosken	30	U	106	351	Forest					
18	2	East Tazenda	280	Ludza	06	U	186	665	Forest					
19	2	East Tazenda	280	Ludza	16	U	261	555	Forest					
20	2	East Tazenda	280	Ludza	38	U	132	473	Forest					

65. *Probabilités de sélection et pondérations d'échantillonnage.* Les probabilités de sélection de la première phase P(1) sont faciles à calculer sur la feuille de calcul de l'échantillon. On multiplie le nombre de ménages de l'UPE sélectionnée par le nombre d'UPE sélectionnées dans chaque strate (colonnes G et K de la figure XV.9 ci-après) et on divise le résultat par le nombre total de ménages de la strate (colonne J). Ces calculs sont exprimés par la formule = $K2*G2/J2$ inscrite dans la case L2, et figurent dans la colonne L.

66. Bien entendu, les probabilités de sélection aux phases suivantes sont fonction des caractéristiques particulières de l'échantillon. Nous illustrerons les calculs pour un échantillonnage à deux phases avec un nombre fixe de ménages sélectionnés selon la même probabilité dans chaque UPE de la seconde phase. En fait, cette conception de l'échantillon est l'une des plus couramment utilisées. Le nombre de ménages par UPE sélectionnée dans la seconde phase peut varier selon les strates, mais dans le pays hypothétique de notre exemple, nous supposons qu'il y a 12 ménages par ZE dans toutes les strates.

67. Cette phase de l'échantillonnage nécessite généralement l'établissement d'une liste des ménages de chaque UPE sélectionnée. Ces listes n'ont pas à être informatisées car la sélection des ménages à visiter pour l'enquête peut se faire à la main à partir des listes sur papier. Toutefois, il y a de nombreux avantages à avoir ces listes sur fichiers informatiques (par exemple, si les UPE sélectionnées lors de la première phase constituent un échantillon-maître qui doit servir à diverses enquêtes ou à divers cycles d'une même enquête).

Figure XV.9.

Calcul des probabilités de la sélection de la première phase

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Province	Province Name	Ward	Ward Name	CEA	Urban/Rural	Households	Population	Stratum	Number of HHs in the Stratum	Number of PSUs in the Stratum	P(1)		
2	1	West Tazenda	207	Macondo	17	U	53	312	Forest	110,388	88	0.04225		
3	1	West Tazenda	226	Balayan	01	U	125	483	Forest	110,388	88	0.09965		
4	1	West Tazenda	226	Balayan	53	U	69	394	Forest	110,388	88	0.05501		
5	1	West Tazenda	226	Balayan	90	U	43	192	Forest	110,388	88	0.03428		
6	1	West Tazenda	242	Haliyil	52	U	48	279	Forest	110,388	88	0.03827		
7	1	West Tazenda	255	Gronau	15	U	52	333	Forest	110,388	88	0.04145		
8	1	West Tazenda	259	Pazar	04	U	79	395	Forest	110,388	88	0.06298		
9	1	West Tazenda	401	Tolbo	21	U	84	361	Forest	110,388	88	0.06696		
10	1	West Tazenda	401	Tolbo	38	U	130	463	Forest	110,388	88	0.10363		
11	2	East Tazenda	267	Xanadu	06	U	125	511	Forest	110,388	88	0.09965		
12	2	East Tazenda	267	Xanadu	25	U	105	424	Forest	110,388	88	0.08370		
13	2	East Tazenda	270	Quetta	02	U	166	580	Forest	110,388	88	0.13233		
14	2	East Tazenda	270	Quetta	21	U	138	407	Forest	110,388	88	0.11001		
15	2	East Tazenda	270	Quetta	45	U	177	485	Forest	110,388	88	0.14110		
16	2	East Tazenda	275	Mosken	10	U	150	368	Forest	110,388	88	0.11958		
17	2	East Tazenda	275	Mosken	30	U	106	351	Forest	110,388	88	0.08450		
18	2	East Tazenda	280	Ludza	06	U	186	665	Forest	110,388	88	0.14828		
19	2	East Tazenda	280	Ludza	16	U	261	555	Forest	110,388	88	0.20807		

68. En général, le nombre de ménages que l'on trouve dans chaque UPE sélectionnée au moment de l'établissement des listes diffère du « nombre de ménages » enregistrés à l'origine lors du recensement dans le cadre d'échantillonnage de la première phase. **Il convient d'ajouter une colonne sur la feuille de calcul de l'échantillon pour inscrire le nombre de ménages énumérés.** Si ces listes sont informatisées, cette colonne peut être remplie par le programme (par exemple, à l'aide des macros Excel). Sinon, les responsables de la gestion de l'enquête doivent remplir cette colonne en priorité, dans le cadre de l'établissement des listes de ménages. Sur la figure XV.10, le « nombre de ménages » du cadre et le « nombre de ménages énumérés » figurent respectivement dans les colonnes G et M.

69. Lorsque le travail de terrain et la gestion des données sont terminés, il convient d'ajouter de nouvelles colonnes à la feuille de calcul de l'échantillon afin d'enregistrer, **pour chaque UPE**, le nombre des ménages pour lesquels des informations utiles ont été enregistrées dans les ensembles de données de l'enquête, ainsi que le nombre des ménages pour lesquels, pour des raisons diverses, on ne dispose pas d'informations. Les raisons typiques pour ajouter une colonne « questionnaire inutile » pour cause de non-réponse sont examinées longuement ailleurs dans la présente publication (voir, par exemple, chapitre VIII et section F du chapitre XXII pour les cas de refus, de logement vacant, etc.). Une telle colonne peut également avoir à être ajoutée lorsque l'enquête ne peut intégrer de contrôles informatiques de qualité aux

opérations sur le terrain. Tel est malheureusement le résultat courant de l'application de techniques d'entrée centralisée des données.

Figure XV.10.
Résultats de l'établissement des listes de ménages

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Province	Province Name	Ward	Ward Name	CEN	Urban/Rural	Households	Population	Stratum	Number of HHs in the Stratum	Number of PSUs in the Stratum	P(1)	Number of HHs Listed	
2	1	West Tazenda	207	Macondo	17	U	53	312	Forest	110,388	88	0.04225	58	
3	1	West Tazenda	226	Balayan	01	U	125	483	Forest	110,388	88	0.09965	153	
4	1	West Tazenda	226	Balayan	53	U	69	394	Forest	110,388	88	0.05501	69	
5	1	West Tazenda	226	Balayan	90	U	43	192	Forest	110,388	88	0.03428	44	
6	1	West Tazenda	242	Haliyal	52	U	48	279	Forest	110,388	88	0.03827	62	
7	1	West Tazenda	255	Gronau	15	U	52	333	Forest	110,388	88	0.04145	46	
8	1	West Tazenda	259	Pazar	04	U	79	395	Forest	110,388	88	0.06298	74	
9	1	West Tazenda	401	Tolbo	21	U	84	361	Forest	110,388	88	0.06696	95	
10	1	West Tazenda	401	Tolbo	38	U	130	463	Forest	110,388	88	0.10363	90	
11	2	East Tazenda	267	Xanadu	06	U	125	511	Forest	110,388	88	0.09965	117	
12	2	East Tazenda	267	Xanadu	25	U	105	424	Forest	110,388	88	0.08370	101	
13	2	East Tazenda	270	Quetta	02	U	166	580	Forest	110,388	88	0.13233	174	
14	2	East Tazenda	270	Quetta	21	U	138	407	Forest	110,388	88	0.11001	138	
15	2	East Tazenda	270	Quetta	45	U	177	485	Forest	110,388	88	0.14110	182	
16	2	East Tazenda	275	Mosken	10	U	150	368	Forest	110,388	88	0.11958	150	
17	2	East Tazenda	275	Mosken	30	U	106	351	Forest	110,388	88	0.08450	132	
18	2	East Tazenda	280	Ludza	06	U	186	665	Forest	110,388	88	0.14828	191	
19	2	East Tazenda	280	Ludza	16	U	261	555	Forest	110,388	88	0.20807	285	

70. Pour continuer avec l'exemple présenté à la figure XV.11 ci-dessous, nous simplifierons la situation en supposant que deux colonnes supplémentaires sont ajoutées à la feuille de calcul, pour indiquer le « nombre de ménages figurant dans les ensembles de données » et le total des cas de « non-réponse ».

71. Bien qu'il n'y ait pas de modèle universellement accepté pour les cas de non-réponse, on suppose généralement que les ménages « utiles » des ensembles de données retenus constituent en fait un échantillon de même probabilité que l'ensemble des ménages énumérés dans leurs UPE respectives (voir l'analyse détaillée aux chapitres II et VIII). Selon cette hypothèse, on peut calculer la probabilité P(2) que chacun de ces ménages soit choisi dans la deuxième phase simplement en divisant le nombre de ménages utiles par le nombre de ménages énumérés. La probabilité totale de sélection de chaque ménage de l'UPE est le produit $P(1) \cdot P(2)$ et la pondération est l'inverse de cette probabilité.

72. Ces formules peuvent facilement s'appliquer au tableur (voir figure XV.12). Écrire la formule $=N2/M2$ dans la case P2, la formule $=L2 \cdot P2$ dans la case Q2 et la formule $=1/Q2$ dans la case R2, puis les appliquer respectivement à toutes les cases des colonnes P, Q et R.

Figure XV.11.
Non-réponse

Province	Province Name	Ward	Ward Name	CEA	Urban/Rural	Households	Population	District	Number of Hqs in the District	Number of FDU's in the District	P(1)	Number of Hqs Listed	Number of Hqs in the dataset	Non-response
1	West Tazenda	207	Macondo	17	U	53	312	Forest	110,388	88	0.04225	58	11	1
2	West Tazenda	226	Bakayan	21	U	128	483	Forest	110,388	88	0.09968	153	12	0
3	West Tazenda	226	Bakayan	33	U	69	394	Forest	110,388	88	0.08501	69	12	0
4	West Tazenda	226	Bakayan	30	U	43	192	Forest	110,388	88	0.03428	44	11	1
5	West Tazenda	242	Nalalyal	32	U	48	279	Forest	110,388	88	0.03527	62	8	4
6	West Tazenda	251	Gronku	15	U	52	333	Forest	110,388	88	0.04148	48	11	1
7	West Tazenda	258	Pecar	04	U	79	385	Forest	110,388	88	0.06298	74	11	1
8	West Tazenda	401	Tobo	21	U	84	361	Forest	110,388	88	0.06694	95	12	0
9	West Tazenda	401	Tobo	28	U	130	463	Forest	110,388	88	0.10363	90	8	4
10	East Tazenda	267	Xenadu	06	U	125	511	Forest	110,388	88	0.09965	117	12	0
11	East Tazenda	267	Xenadu	25	U	105	424	Forest	110,388	88	0.06575	101	12	0
12	East Tazenda	270	Qurtta	02	U	166	580	Forest	110,388	88	0.13233	174	11	1
13	East Tazenda	270	Qurtta	21	U	138	457	Forest	110,388	88	0.115001	138	10	2
14	East Tazenda	270	Qurtta	45	U	177	488	Forest	110,388	88	0.14110	162	12	0
15	East Tazenda	278	Mookien	10	U	190	368	Forest	110,388	88	0.11868	190	12	0
16	East Tazenda	278	Mookien	30	U	108	351	Forest	110,388	88	0.09450	122	11	1
17	East Tazenda	280	Leitza	06	U	188	685	Forest	110,388	88	0.14628	191	11	1
18	East Tazenda	280	Leitza	16	U	261	550	Forest	110,388	88	0.20807	260	11	1

Figure XV.12
Calcul des probabilités et des pondérations de la deuxième phase

Province	Province Name	Ward	Ward Name	CEA	Urban/Rural	Households	Population	District	Number of Hqs in the District	Number of FDU's in the District	P(1)	Number of Hqs Listed	Number of Hqs in the dataset	Non-response	P(2)	P(1)*P(2)	Weight
1	West Tazenda	207	Macondo	17	U	53	312	Forest	110,388	88	0.04225	58	11	1	0.18986	0.00801	124.80
2	West Tazenda	226	Bakayan	21	U	128	483	Forest	110,388	88	0.09968	153	12	0	0.87943	0.08782	127.99
3	West Tazenda	226	Bakayan	33	U	69	394	Forest	110,388	88	0.08501	69	12	0	0.17391	0.00987	104.53
4	West Tazenda	226	Bakayan	30	U	43	192	Forest	110,388	88	0.03428	44	11	1	0.25000	0.00857	118.89
5	West Tazenda	242	Nalalyal	32	U	48	279	Forest	110,388	88	0.03527	62	8	4	0.12903	0.00494	202.53
6	West Tazenda	251	Gronku	15	U	52	333	Forest	110,388	88	0.04148	48	11	1	0.23913	0.00991	100.88
7	West Tazenda	258	Pecar	04	U	79	385	Forest	110,388	88	0.06298	74	11	1	0.14888	0.00938	106.82
8	West Tazenda	401	Tobo	21	U	84	361	Forest	110,388	88	0.06694	95	12	0	0.12632	0.00846	118.22
9	West Tazenda	401	Tobo	28	U	130	463	Forest	110,388	88	0.10363	90	8	4	0.58889	0.00821	108.55
10	East Tazenda	267	Xenadu	06	U	125	511	Forest	110,388	88	0.09965	117	12	0	0.10256	0.01022	97.84
11	East Tazenda	267	Xenadu	25	U	105	424	Forest	110,388	88	0.06575	101	12	0	0.11581	0.00985	106.55
12	East Tazenda	270	Qurtta	02	U	166	580	Forest	110,388	88	0.13233	174	11	1	0.06322	0.00837	118.93
13	East Tazenda	270	Qurtta	21	U	138	457	Forest	110,388	88	0.115001	138	10	2	0.87348	0.00787	125.44
14	East Tazenda	270	Qurtta	45	U	177	488	Forest	110,388	88	0.14110	162	12	0	0.06883	0.00930	107.46
15	East Tazenda	278	Mookien	10	U	190	368	Forest	110,388	88	0.11868	190	12	0	0.08000	0.00987	104.53
16	East Tazenda	278	Mookien	30	U	108	351	Forest	110,388	88	0.09450	122	11	1	0.08333	0.00794	142.01
17	East Tazenda	280	Leitza	06	U	188	685	Forest	110,388	88	0.14628	191	11	1	0.08789	0.00894	117.10
18	East Tazenda	280	Leitza	16	U	261	550	Forest	110,388	88	0.20807	260	11	1	0.02880	0.00903	124.52

73. Les pondérations fondées sur les probabilités calculées de cette manière s'appliquent à tous les ménages de chaque UPE. Certains spécialistes des enquêtes peuvent utiliser des techniques de « post-stratification » pour ajuster davantage ces pondérations de manière que les estimations de l'enquête concordent avec les répartitions de population connues (telles que les répartitions par âge et par sexe ou les chiffres de consommation totale obtenus de sources extérieures à l'enquête elle-même). Ces ajustements sont appliqués directement aux ensembles de données de l'enquête à l'aide d'un logiciel spécialisé, et non pas aux tableurs, et se font généralement individuellement pour chaque ménage ou chaque personne plutôt que pour l'UPE.

H. Résumé des recommandations

74. Dans ce chapitre, nous nous sommes efforcés de montrer l'utilité d'incorporer les critères de gestion des données à chaque phase de l'enquête, plutôt que de considérer ces critères comme ne s'appliquant qu'aux dernières phases d'analyse. L'un des exemples les plus clairs nous est donné par les enquêtes de l'Étude de la mesure des niveaux de vie, qui ont englobé la conception des questionnaires, ainsi que la préparation et l'exécution des opérations sur le terrain, et ont traité l'entrée et le traitement des données de manière à permettre une bonne gestion de ces données avant même leur collecte. Ces principes directeurs ont été au centre de ce chapitre, et même s'ils affichent des caractéristiques différentes selon l'application qui en est faite dans un pays donné, ces principes peuvent encore se résumer et se définir dans les termes suivants :

- a) La gestion des données de l'enquête commence à la conception du questionnaire et, à ce stade, traite de :
 - i) La bonne identification des unités statistiques. Il est recommandé d'utiliser des numéros de série simples ou améliorés de trois ou quatre chiffres pour les UPE de l'enquête, puis un numéro de série de deux chiffres pour chaque ménage de l'UPE, plus un numéro d'identification de chaque unité secondaire au sein du ménage;
 - ii) Redondances délibérées. Le questionnaire doit inclure des redondances délibérées, destinées à signaler les erreurs de l'enquêteur ou les erreurs d'entrée des données. Par exemple, il peut prévoir une ligne de totaux ou ajouter un chiffre de contrôle aux codes des variables importantes.
- b) Lors des opérations sur le terrain, il faut tenir compte de ce qui suit :
 - i) Les stratégies opérationnelles pour l'entrée et l'édition des données. Il est recommandé que les pays considèrent avec soin la possibilité d'entrer toutes les données sur le terrain. Cette tâche peut être confiée à un préposé à l'entrée des données travaillant à un lieu fixe autre que celui où se trouvent les ménages, par un opérateur s'associant au reste de l'équipe chargée de l'enquête et entrant les données directement sur un ordinateur portable au domicile de chaque ménage ou selon la méthode non encore suffisamment étudiée de l'entrevue sans papier à l'aide d'un ordinateur de poche (quoique cela nécessite davantage de recherche). Le fait d'entrer les données sur le terrain plutôt qu'en un lieu central contribue grandement à assurer la qualité et l'homogénéité des données;
 - ii) Les critères de contrôle de qualité. Les données entrées sur les questionnaires doivent être soumises à cinq mécanismes de contrôle différents : vérifications de l'ordre de grandeur; vérifications par rapport à

- des tableaux de référence; vérifications des questions à sauter; vérifications de la cohérence et vérifications typographiques;
- iii) La technologie d'entrée des données. D'après une étude de 1995 de la Banque mondiale, les deux plates-formes acceptables pour l'entrée et l'édition des données d'enquêtes complexes sur les ménages étaient celles de l'enquête LSMS conçue par les services de la Banque mondiale et le programme IMPS d'United States Bureau of the Census. Les versions actualisées de ces plates-formes sont respectivement LSD-2000 et CSPro. Si l'on tient compte des compétences et d'autres facteurs influant sur la situation propre à chaque pays, il y a peu de directives à prendre en compte lors de la conception des outils d'entrée et d'édition des données : à quelques exceptions près, les écrans d'ordinateur devraient ressembler aux sections correspondantes du questionnaire; les programmes d'entrée des données devraient pouvoir discerner les situations impossibles ou improbables et les signaler; le langage et les expressions servant à signaler le type d'erreur doivent être simples et faciles à comprendre;
 - iv) *L'organisation et la diffusion des ensembles de données de l'enquête.* Pour ces usages, les fichiers plats ne conviennent pas, car ils ne peuvent traiter convenablement les unités statistiques secondaires (personnes, cultures, articles de consommation, etc.) au sein du ménage. Une structure offrant différents types d'enregistrement pour chaque type d'unité statistique est jugée préférable.
- c) Enfin, la gestion des données peut aussi s'avérer importante pour l'application du protocole d'échantillonnage, en le guidant à travers ses principales étapes : organisation du cadre d'échantillonnage de la première phase, créé généralement à partir de l'ensemble de zones d'énumération (ZE) du dernier recensement; sélection d'unités primaires d'échantillonnage selon une probabilité proportionnelle à la taille, mesurée au nombre de ménages ou de logements ou à la taille de la population; et calcul des probabilités de sélection et des pondérations d'échantillonnage correspondantes.

RÉFÉRENCES

- Ainsworth, M., et J. Muñoz (1986). *The Côte d'Ivoire Living Standards Survey: Design and Implementation*. Living Standards Measurement Study Working Paper, No. 26. Washington, D.C., Banque mondiale.
- Blaizeau, D. (1998). Seven expenditure surveys in the West African Economic and Monetary Union. Dans *Proceedings of the Joint International Association of Survey Statisticians/International Association for Official Statistics (IASS/IAOS) Conference on Statistics for Economic and Social Development*. Aguascalientes, Mexique, International Statistical Institute.
- _____, et J. L. Dubois (1990). *Connaître les conditions de vie des ménages dans les pays en développement*. Paris, Documentation française.
- Blaizeau, D, et J. Muñoz (1998). *LSD-2000. Logiciel de saisie des données : pour saisir les données d'une enquête complexe*. Paris, Institut national de la statistique et des études économiques.
- Grosh, M. et J. Muñoz (1996). *A Manual for Planning and Implementing the Living Standards Measurement Study Survey*, Living Standards Measurement Study Working Paper, No. 126. Washington, D.C., Banque mondiale.

Muñoz, J. (1989). Data management of complex socioeconomic surveys: from questionnaire design to data analysis. Dans *Proceedings of the 47th Session of the International Statistical Institute*. Paris, Institut international de statistique.

_____ (1996). Cómo mejorar la calidad de la información: opciones para mejorar la organización del trabajo de campo, el sistema de entrada de datos, el análisis de consistencia y el manejo de la base de datos. dans *Reunión de Iniciación del Programa para el Mejoramiento de las Encuestas de Condiciones de Vida en América Latina y El Caribe*. Asunción, Banque interaméricaine de développement.

_____ (1998). Budget-Consumption Surveys: New Challenges and Outlook. Dans *Proceedings of the Joint International Association of Survey Statisticians/International Association for Official Statistics (IASS/IAOS) Conference on Statistics for Economic and Social Development*. Aguascalientes, Mexique, Institut international de statistique.

United States Bureau of the Census. CPro Census and Survey Processing System. Peut être consulté sur <http://www.census.gov/ipc/www/cspro/>.

Chapitre XVI

Présentation de statistiques descriptives simples extraites des données d'une enquête sur les ménages

PAUL GLEWWE

Department of Applied Economics
University of Minnesota
Saint Paul, Minnesota, États-Unis d'Amérique

MICHAEL LEVIN

United States Bureau of the Census
Washington, D.C.,
États-Unis d'Amérique

RÉSUMÉ

Le présent chapitre donne des directives générales pour le calcul et la présentation de statistiques descriptives de base pour les données d'enquêtes sur les ménages. L'analyse est élémentaire en ce sens qu'elle consiste à présenter des tableaux relativement simples et faciles à comprendre par un large public. Ce chapitre donne également des conseils sur la façon de présenter les tableaux et les graphiques dans un rapport général destiné à une large diffusion.

Termes clés : statistiques descriptives, tableaux, graphiques, résumé statistique, diffusion.

A. Introduction

1. La valeur vraie des données des enquêtes sur les ménages ne s'obtient que lorsque ces données sont analysées. L'analyse des données peut aller des statistiques sommaires très simples à des analyses multivariates extrêmement complexes. Le présent chapitre sert d'introduction aux quatre chapitres suivants et, à ce titre, portera sur les questions élémentaires et les méthodes relativement simples. Les quatre chapitres suivants traiteront de questions plus complexes.

2. La plupart des données des enquêtes sur les ménages peuvent être utilisées de diverses façons pour illustrer les phénomènes qui sont le principal centre d'intérêt de l'enquête. Dans un sens, le point de départ de l'analyse des données se situe dans les statistiques descriptives de base telles que les tableaux des moyennes et des fréquences des principales variables étudiées. Toutefois, le point de départ le plus élémentaire de l'analyse réside dans les questions auxquelles les données recueillies sont destinées à répondre. Autrement dit, dans presque toutes les enquêtes sur les ménages, la première tâche est de définir les objectifs de l'enquête, afin que les données recueillies soient de nature à permettre la réalisation de ces objectifs. Cela implique que la conception de l'enquête et la planification en vue de l'analyse des données aient lieu simultanément avant la collecte des données. Cela est expliqué de façon plus détaillée au chapitre III. Le présent chapitre portera sur de nombreux aspects pratiques de l'analyse des don-

nées, à supposer qu'une stratégie judicieuse ait déjà été conçue pour l'analyse des données conformément aux conseils donnés au chapitre III.

3. Ce chapitre s'articule de la façon suivante : la section B passe en revue les types de variables et les statistiques descriptives simples; la section C contient des conseils d'ordre général sur la façon de préparer et de présenter les statistiques de base tirées des données de l'enquête; et la section D formule des recommandations sur la façon de préparer un rapport général (souvent appelé résumé statistique) qui présente les résultats essentiels de l'enquête sur les ménages à une large audience. La brève section finale offre certaines conclusions.

B. Variables et statistiques descriptives

4. Beaucoup d'enquêtes sur les ménages servent à recueillir des données sur un sujet ou un thème particulier, tandis que d'autres en recueillent sur une variété de sujets. Dans un cas comme dans l'autre, les données recueillies peuvent être considérées comme un ensemble de variables dont certaines présentent un intérêt à elles seules tandis que d'autres sont surtout intéressantes quand on les compare à d'autres. Un grand nombre de ces variables varient au niveau du ménage; tel est le cas du type de logement. D'autres peuvent varier au niveau de l'individu, comme l'âge et la situation de famille. Certaines enquêtes visent à recueillir des données qui ne varient qu'au niveau de la communauté, comme le prix de divers biens vendus sur le marché local¹.

5. La première étape de toute analyse de données est de produire un ensemble de données contenant toutes les variables à étudier. Les analystes peuvent alors calculer des statistiques descriptives de base qui permettent aux variables de s'expliquer d'elles-mêmes. Il existe un nombre relativement restreint de façons de procéder. La présente section explique comment cela se fait. Elle commence par un bref examen des différents types de variables et de statistiques descriptives, puis traite des méthodes de présentation des données sur une seule variable, sur deux variables et sur trois variables ou plus.

1. Types de variables

6. Les enquêtes sur les ménages servent à recueillir des données sur deux types de variables : les variables « catégoriques » et les variables « numériques ». Les variables catégoriques sont des caractéristiques qui ne sont pas définies par des chiffres mais par des catégories ou des types. Comme exemples de variables catégoriques, on peut citer les caractéristiques des logements (type de sol, matériau utilisé pour les murs, type de toilettes, etc.), et les caractéristiques personnelles, telles que le groupe ethnique, la situation de famille et la profession. Dans la pratique, on peut affecter des codes numériques à ces caractéristiques, et désigner un groupe ethnique comme « code 1 », un autre comme « code 2 » et ainsi de suite, mais ce n'est qu'une convention arbitraire. En revanche, les variables numériques sont des chiffres de par leur nature même. Le nombre de pièces d'un logement, l'étendue des terres détenues ou le revenu d'un membre particulier du ménage sont par définition des variables numériques. Tout au long de ce chapitre, les différents types possibles des variables catégoriques seront définis comme « catégories » et les différentes valeurs des variables numériques seront définies comme « valeurs ».

7. Lorsque l'on présente des données pour l'un ou l'autre de ces types de variables, il est utile de faire une autre distinction concernant le nombre de catégories ou de valeurs que peut prendre la variable en question. Si ce nombre est faible, par exemple, inférieur à 10, il est alors commode (et instructif) d'afficher toutes les informations

¹ Dans la plupart des enquêtes sur les ménages, le ménage est défini comme un groupe d'individus qui : a) vivent dans le même logement; b) prennent au moins un repas par jour ensemble ; et c) mettent en commun leurs revenus et autres ressources pour l'achat de biens et de services. Certaines enquêtes sur les ménages modifient cette définition au gré des circonstances locales, mais cette question déborde le cadre de ce chapitre. La « communauté » est plus difficile à définir, mais aux fins du présent chapitre, on peut la considérer comme un ensemble de ménages vivant dans le même village, la même ville ou le même quartier. [Voir Frankenberg (2000) pour un examen détaillé de la définition de la « communauté ».]

sur la distribution de la variable. En revanche, si ce nombre est élevé, par exemple, supérieur à 10, il est généralement préférable de n'afficher que des statistiques globales ou sommaires concernant cette distribution. Un exemple éclairera le lecteur. Dans un pays, la population peut être composée d'un petit nombre de groupes ethniques, quatre, par exemple. Pour ce pays, il est relativement facile de présenter sur un simple tableau ou graphique le pourcentage des ménages sélectionnés appartenant à chaque groupe. Par contre, dans un autre pays, il peut y avoir des centaines de groupes ethniques. Il serait très laborieux de présenter le pourcentage des ménages sélectionnés appartenant à chacun de ces groupes si ces derniers sont, par exemple, au nombre de 400. Dans la plupart des cas, il serait plus simple et suffisant de grouper plusieurs groupes ethniques différents en un petit nombre de catégories plus larges et d'afficher le pourcentage de ménages entrant dans chacune de ces catégories.

8. L'exemple ci-dessus utilisait une variable catégorique, le groupe ethnique, mais il s'applique aussi aux variables numériques. Certaines variables numériques, telles que le nombre de jours où une personne a été malade la semaine précédente, ne peuvent avoir qu'un très faible nombre de valeurs et leur distribution peut être représentée intégralement sur un tableau ou graphique simple. En revanche, de nombreuses autres variables numériques, telles que le nombre d'animaux de ferme détenus, peuvent prendre un grand nombre de valeurs différentes, et il est préférable de ne présenter que des statistiques sommaires de leur distribution. La principale différence entre le traitement des variables catégoriques et numériques découle de la façon de les grouper lorsque le nombre des valeurs/catégories possibles est élevé. Pour les variables catégoriques, une fois qu'il a été décidé de ne pas montrer intégralement leur distribution, on n'a pas d'autre choix que de les grouper par grandes catégories. Pour les variables numériques, il est possible de les grouper en grandes catégories, mais on peut aussi afficher des statistiques sommaires, telles que la moyenne, l'écart type, voir les valeurs minimum et maximum. La sous-section suivante passe brièvement en revue les statistiques descriptives les plus courantes.

2. Statistiques descriptives simples

9. Les tableaux et graphiques peuvent fournir certaines informations essentielles sur les variables à étudier en utilisant des statistiques descriptives simples telles que, notamment, les distributions en pourcentage, les valeurs médianes, les moyennes et les écarts types. Cette sous-section examine ces statistiques simples, en donnant des exemples tirés de l'enquête sur les ménages de Saipan, qui appartient au Commonwealth des îles Mariannes du Nord et des Samoa américaines.

10. *Distributions en pourcentage* : Les enquêtes sur les ménages recueillent rarement des données pour 100, 1 000 ou 10 000 personnes ou ménages. Supposons que l'on ait des données sur les catégories d'une variable catégorique, telles que les nombres de membres d'une population qui sont respectivement du sexe masculin et du sexe féminin, ou des données sur une variable numérique telle que l'âge en années des membres de cette même population. La présentation des chiffres observés dans chaque catégorie est généralement moins utile que celle du pourcentage des cas observés par catégorie. On peut s'en rendre compte en regardant les chiffres des trois premières colonnes du tableau XVI.1. La plupart des utilisateurs auraient plus de mal à interpréter ces chiffres s'ils n'étaient pas accompagnés des distributions en pourcentage. Les trois dernières colonnes du tableau XVI.1 sont beaucoup plus faciles à comprendre si l'on s'intéresse au pourcentage d'hommes et de femmes par groupe d'âge. C'est ce que montre le tableau XVI.2. (On peut également faire les totaux des pourcentages par groupe d'âge et par sexe, mais cela est généralement moins intéressant.) Ces deux

tableaux montrent que les distributions en pourcentage peuvent être présentées aussi bien pour les variables catégoriques que pour les variables numériques.

Tableau XVI.1

**Distribution de la population par âge et par sexe, Saipan,
Commonwealth des îles Mariannes du Nord, avril 2002 (pourcentages par rangée)**

Groupe d'âge, en années	Chiffres			Pourcentages par rangée		
	Total	Hommes	Femmes	Total	Hommes	Femmes
Total de personnes	67 011	29 668	37 343	100,0	44,3	55,7
Moins de 15 ans	16 915	8 703	8 212	100,0	51,5	48,5
15 à 29 ans	18 950	5 765	13 184	100,0	30,4	69,6
30 à 44 ans	20 803	9 654	11 149	100,0	46,4	53,6
45 à 59 ans	8 105	4 458	3 648	100,0	55,0	45,0
60 ans et plus	2 239	1 088	1 150	100,0	48,6	51,4

Source : Dixième cycle de l'Enquête sur la main-d'œuvre au Commonwealth des îles Mariannes du Nord.

Note : Les données sont tirées d'un échantillon de 10 % des ménages et de la totalité des personnes vivant en communauté.

Tableau XVI.2

**Distribution de la population par âge et par sexe, Saipan,
Commonwealth des îles Mariannes du Nord, avril 2002 (pourcentages par colonne)**

Groupe d'âge, en années	Chiffres			Pourcentages par colonne		
	Total	Hommes	Femmes	Total	Hommes	Femmes
Nombre total de personnes	67 011	29 668	37 343	100,0	100,0	100,0
Moins de 15 ans	16 915	8 703	8 212	25,2	29,3	22,0
5 à 29 ans	18 950	5 765	13 184	28,3	19,4	35,3
30 à 44 ans	20 803	9 654	11 149	31,0	32,5	29,9
45 à 59 ans	8 105	4 458	3 648	12,1	15,0	9,8
60 ans et plus	2 239	1 088	1 150	3,3	3,7	3,1

Source : Dixième cycle de l'Enquête sur la main-d'œuvre au Commonwealth des îles Mariannes du Nord.

Note : Les données sont tirées d'un échantillon de 10 % des ménages et de la totalité des personnes vivant en communauté.

11. Le tableau XVI.1 montre clairement que la distribution par sexe varie selon les groupes d'âge. Cela dénote un phénomène que ne montrent pas les tableaux XVI. 1 et XVI.2, à savoir que Saipan a de nombreux travailleurs immigrés, surtout des femmes, qui travaillent dans les usines de vêtements. Bien que Saipan compte légèrement plus d'éléments masculins que féminins aux jeunes âges, le groupe d'âge suivant, les personnes âgées de 15 à 29 ans, ne compte que 30 hommes pour 70 femmes. Il y a plus de femmes que d'hommes âgés de 30 à 44 ans, ce qui est conforme au fait que la plupart des travailleurs de l'industrie du vêtement sont des femmes âgées de 20 à 40 ans. Dans le groupe d'âge suivant, celui des personnes âgées de 45 à 59 ans, il y a plus d'hommes que de femmes. Les colonnes des pourcentages du tableau XVI.2 montrent que le plus fort groupe d'âge pour les hommes est celui des hommes âgés de 30 à 44 ans, tandis que le plus fort groupe d'âge pour les femmes est celui des femmes âgées de 15 à 29 ans, qui correspond aux âges où les femmes ont le plus de chances de travailler dans l'industrie du vêtement.

12. *Valeurs médianes.* Les deux mesures statistiques les plus courantes pour les variables numériques sont les moyennes et les valeurs médianes. (Par définition, les variables catégoriques ne sont pas numériques et ne se prêtent donc pas au calcul de valeurs moyennes ou médianes.) La valeur médiane est le point milieu d'une distribution, tandis que la valeur moyenne est la moyenne arithmétique des valeurs. On utilise souvent la valeur médiane pour les variables telles que l'âge et le revenu, car elle est moins sensible aux phénomènes extérieurs. Comme exemple extrême, supposons qu'il y a 99 personnes dans une enquête avec des revenus compris entre \$8 000 et \$12 000 par an, et que ces personnes sont symétriquement distribuées autour de \$10 000. Dans ce cas, la moyenne et la médiane seraient toutes deux de \$10 000. À présent, supposons que l'enquête englobe une personne de plus avec un revenu annuel de \$500 000; dès lors, la moyenne serait d'environ \$15 000 tandis que la médiane resterait d'environ \$10 000. Dans beaucoup de cas, les rapports sur les variables de revenu indiquent à la fois la valeur moyenne et la valeur médiane.

13. Pour en revenir aux données de Saipan, l'âge médian de la population de Saipan était de 28,5 ans en avril 2002, ce qui veut dire que la moitié de la population avait plus de 28,5 ans et la moitié avait moins que cet âge. L'âge médian des femmes était inférieur à celui des hommes (27,6 ans contre 30,5), en raison du grand nombre de jeunes immigrantes travaillant dans les usines de vêtements.

14. *Moyennes et écarts types.* Comme on l'a indiqué précédemment, la valeur moyenne est la moyenne arithmétique d'une variable numérique. On calcule souvent la moyenne pour le nombre d'enfants nés (d'une femme), le revenu et d'autres variables numériques. L'écart type mesure la distance moyenne d'une variable numérique à la moyenne de cette variable, et offre donc une indication de la dispersion dans la distribution d'une variable numérique.

15. Le tableau XVI.3 indique les valeurs médianes et moyennes du revenu annuel telles qu'elles ressortent de l'Enquête de 1995 sur les ménages des Samoa américaines. Cette enquête portait sur un échantillon aléatoire de 20 % de tous les ménages du territoire. Le fait que le revenu moyen des ménages était plus élevé que le revenu médian n'est pas étonnant car certains ménages gagnaient des salaires sensiblement plus élevés et tiraient un revenu plus élevé d'autres sources. Les immigrants de Tonga sont relativement pauvres, à en juger par leurs faibles revenus moyen et médian, alors que, d'après leurs revenus moyen et médian, les « autres groupes ethniques » semblent relativement plus aisés.

Tableau XVI.3

Statistiques sommaires sur le revenu des ménages par groupe ethnique, Samoa américaines, 1994

Revenu annuel	Total	Samoans	Tongans	Ethniques
Nombre de ménages interrogés	8 367	7 332	244	790
Revenu médian (<i>en dollars des États-Unis</i>)	15 715	15 786	7 215	23 072
Revenu moyen (<i>en dollars des États-Unis</i>)	20 670	20 582	8 547	25 260

Source : Enquête de 1995 sur les ménages des Samoa américaines.

Note : Les données émanent d'un échantillon aléatoire non pondéré de 20 % des ménages.

3. Présentation de statistiques descriptives pour une variable

16. Lorsque l'on présente des statistiques descriptives provenant d'une enquête sur les ménages, le cas le plus simple est celui qui ne porte que sur une variable. La pré-

sente sous-section illustre comment présenter de telles statistiques pour des variables catégoriques et numériques.

17. *Présentation de l'ensemble de la distribution.* Les variables catégoriques ou numériques qui ne prennent qu'un petit nombre de catégories ou de valeurs, disons moins de 10, sont les plus simples à afficher. On peut utiliser un tableau indiquant l'ensemble de la distribution (en pourcentage) de la variable en indiquant la fréquence de chaque catégorie ou de chaque valeur numérique de la variable. Le tableau XVI.4 en offre un exemple, avec les chiffres de fréquence d'un échantillon (non pondéré) et la distribution en pourcentage pour les principales sources d'éclairage des familles vietnamiennes. Beaucoup d'enquêtes sur les ménages nécessitent le recours à des pondérations pour estimer la distribution d'une variable au sein de la population; dans ce cas, la présentation des fréquences brutes de l'échantillon peut prêter à confusion et n'est donc pas recommandée; l'utilisation de pondérations sera examinée à la section C ci-dessous. (L'enquête effectuée au Viet Nam était fondée sur un échantillon autopondéré, de sorte qu'aucune pondération n'était nécessaire.) Enfin, il est également utile de signaler les erreurs types des fréquences estimatives des pourcentages (voir l'analyse détaillée de cette question au chapitre XXI, qui est encore compliquée par l'utilisation de pondérations et d'autres caractéristiques de l'échantillon de l'enquête).

18. Dans certains cas, le nombre de catégories ou de valeurs que peut prendre la variable est élevé, mais le gros de leur distribution se limite à quelques catégories ou valeurs. En pareils cas, il peut être nécessaire d'indiquer la fréquence d'apparition de chaque catégorie ou valeur. On peut épargner la patience du lecteur du tableau en combinant les cas rares en une rubrique générale « autres ». Par exemple, toute catégorie ou valeur apparaissant avec une fréquence de moins de 1 % pourrait aller dans cette rubrique. En vérité, c'est ce qui a été fait au tableau XVI.4, où la rubrique « autres » regroupe les cas rares tels que lampes torches. Dans certains cas, il peut y avoir d'autres groupes naturels. Par exemple, dans beaucoup de pays, les groupes ethniques et religieux peuvent se diviser en un grand nombre de catégories distinctes, mais il peut arriver que ces catégories particulières s'inscrivent dans un nombre beaucoup plus faible de groupes généraux. Souvent, il suffit de n'indiquer de chiffres que pour ces groupes. La principale exception à cette règle a trait aux catégories qui, bien que peu répandues, peuvent présenter un intérêt particulier. En général, ces catégories « rares mais présentant un intérêt particulier » peuvent être signalées séparément, mais il est essentiel d'indiquer les erreurs types pour ces cas, car la précision des estimations est plus faible pour les catégories rares.

Tableau XVI.4

Types d'éclairage utilisés par les familles vietnamiennes, 1992-1993

Type	Nombre de familles	Pourcentage de familles (erreur type)
Électricité	2 333	48,6 (0,7)
Kérosène/pétrole lampant	2 386	49,7 (0,7)
Autres	81	1,7 (0,2)
Total des familles de l'échantillon	4 800	100,0

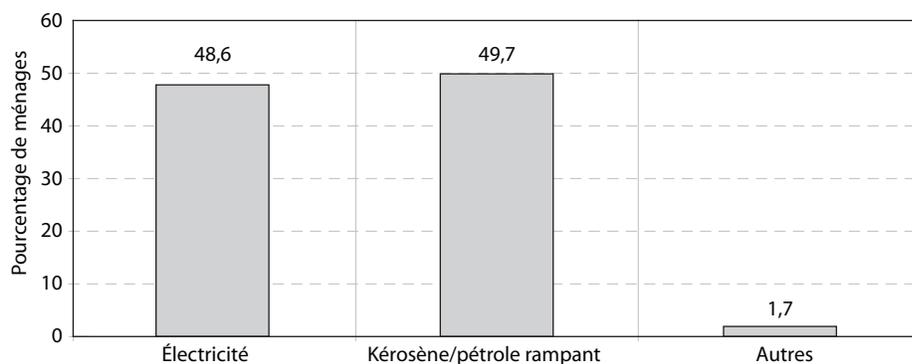
Source : Enquête de 1992-1993 sur les niveaux de vie au Viet Nam.

Note : Les données ne sont pas pondérées.

19. Dans beaucoup de cas, la présentation des données peut être rendue plus intéressante si elle est faite sous forme de graphique ou de carte plutôt que de tableau. Une variable unique qui ne compte qu'un petit nombre de catégories ou de valeurs

peut généralement être représentée sous forme graphique à l'aide d'un graphique en colonnes où la fréquence relative de chaque catégorie ou valeur est indiquée par la hauteur de la colonne. La figure XVI.1 en donne une illustration, utilisant les données présentées au tableau XVI.4. L'autre façon de présenter fréquemment la fréquence relative de catégories ou de valeurs d'une variable est le diagramme circulaire, qui indique la fréquence relative d'une valeur par l'ampleur de la « tranche du gâteau ». La figure XVI.2, qui illustre également les informations données au tableau XVI.4, en offre un exemple. Pour tous conseils sur la façon de concevoir des graphiques, voir Tufte (1983) et Wild et Seber (2000).

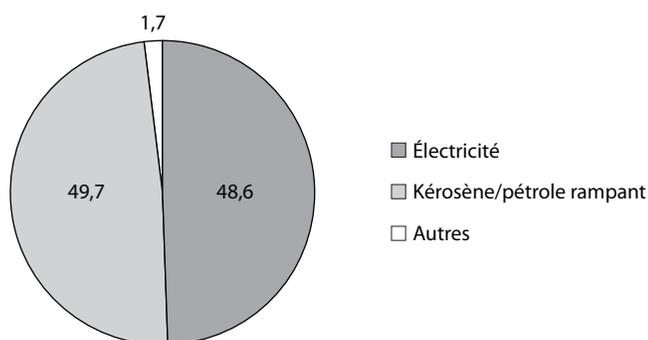
Figure XVI.1
Types d'éclairage utilisés par les familles vietnamiennes, 1992-1993(graphique)
[en pourcentage]



Source : Enquête de 1992-1993 sur les niveaux de vie au Viet Nam.

Note : Taille de l'échantillon : 4 800 familles.

Figure XVI. 2
Types d'éclairage utilisés par les familles vietnamiennes, 1992-1993
(diagramme circulaire) [en pourcentage]



Source : Enquête de 1992-1993 sur les niveaux de vie au Viet Nam.

Note : Taille de l'échantillon : 4 800 familles.

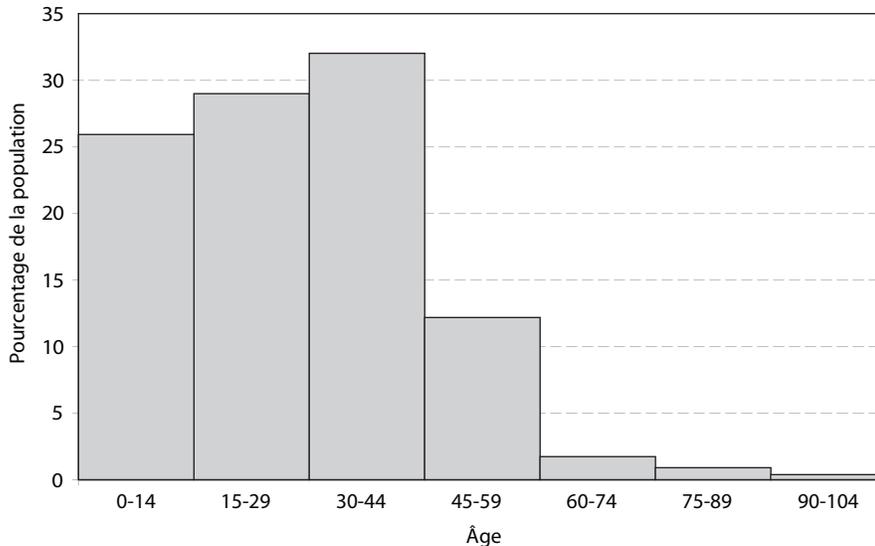
20. *Présentation de variables qui peuvent compter de nombreuses catégories ou valeurs.* Les variables catégoriques comme les variables numériques comptent souvent de nombreuses catégories ou valeurs possibles. Pour les variables catégoriques, la seule façon de présenter des tableaux et graphiques très détaillés est de grouper les catégories en groupes et/ou de combiner toutes les valeurs peu communes en une catégorie

« autres », comme on l'a vu précédemment. Pour les variables numériques, il y a deux possibilités.

21. Premièrement, on peut diviser en un petit nombre de tranches la gamme des nombreuses valeurs d'une variable numérique et afficher les informations de l'une quelconque des façons décrites ci-dessus pour le cas où une variable ne compte qu'un faible nombre de catégories ou de valeurs. Par exemple, c'est ce que l'on a fait pour la variable « âge » aux tableaux XVI.1 et XVI.2. On peut également utiliser cette option pour les graphiques : les informations sur la distribution d'une variable numérique qui peut prendre de nombreuses valeurs peuvent être illustrées par un graphique qui indique la fréquence avec laquelle la variable entre dans un petit nombre de catégories. Un exemple de ce type de graphique est celui de l'histogramme, qui donne une représentation approximative de la densité relative de la variable. Les histogrammes divisent le spectre d'une variable numérique en un nombre relativement faible de zones représentées par des colonnes dont l'ampleur est proportionnelle au pourcentage de l'échantillon qui entre dans la zone correspondante. La figure XVI.3 donne une telle illustration des groupes d'âge du tableau XVI.2. La première colonne correspond au groupe d'âge de 0 à 14 ans, la suivante, au groupe d'âge de 15 à 29 ans, et ainsi de suite². On remarquera qu'à la différence du graphique de la figure XVII.1, il n'y a pas d'intervalles entre les colonnes de l'histogramme. En effet, l'axe des abscisses d'un histogramme décrit l'ensemble des valeurs possibles de la variable, ensemble qui est ininterrompu.

² Cet histogramme divise la population âgée de 60 à 99 ans en trois groupes (60-74 ans, 75-89 ans et 90-104 ans) qui comptent chacun le même nombre d'années (15), comme les groupes de population de moins de 60 ans. Cela permet d'assurer que la surface de chaque colonne de l'histogramme est proportionnelle au pourcentage de la population de chaque groupe d'âge.

Figure XVI.3
Distribution par âge de la population de Saipan, avril 2002 (histogramme)



Source : Dixième cycle de l'Enquête sur la main-d'œuvre au Commonwealth des îles Mariannes du Nord.

22. La seconde possibilité, peut-être plus courante, de présenter une variable numérique qui peut prendre de nombreuses valeurs est de présenter quelques statistiques sommaires de sa distribution, telles que sa moyenne, sa valeur médiane et son écart type. Cela ne peut se faire qu'en présentant ces statistiques sur un tableau; il n'est pas possible de présenter de statistiques sommaires pour une seule variable numérique sur un graphique. Outre la moyenne, la valeur médiane et l'écart type, il est également utile de présenter ses valeurs maximum et minimum, les valeurs des quartiles supé-

rieur et inférieur³, voire une indication de la distorsion. Le tableau XVI.5 en donne un exemple.

Tableau XVI.5
Données récapitulatives sur les dépenses totales des ménages : Viet Nam, 1992-1993
(en milliers de dongs par an)

Moyenne	6 531
Écart type	5 375
Valeur médiane	5 088
Quartile inférieur	3 364
Quartile supérieur	7 900
Valeur la plus faible	235
Valeur la plus grande	100 478

Source : Enquête de 1992-1993 sur les niveaux de vie au Viet Nam.

Note : Taille de l'échantillon : 4 799 800 familles.

³ Le quartile inférieur d'une distribution se termine à la valeur en dessous de laquelle se situent 25 % des observations et le quartile supérieur commence à la valeur au-dessus de laquelle se situent 25 % des observations.

4. Présentation de statistiques descriptives de deux variables.

23. L'examen des relations entre deux variables ou plus éclaire souvent davantage sur le sujet étudié que l'examen d'une seule variable isolée. En même temps, les possibilités d'afficher les données augmentent de façon exponentielle. La présente sous-section décrit diverses méthodes courantes en distinguant les variables selon que leurs catégories et valeurs possibles sont peu nombreuses ou nombreuses.

24. *Deux variables aux nombres de catégories et de valeurs peu élevés.* Le cas le plus simple de présentation de la relation entre deux variables est celui où ces dernières ne peuvent compter qu'un petit nombre de catégories ou de valeurs. Sur le tableau, les catégories ou valeurs d'une variable peuvent être les colonnes et celles de l'autre variable peuvent être les rangées. L'exemple du tableau XVI.6 présente l'utilisation de différents types de prestataires de services de santé dans les zones urbaines et rurales du Viet Nam. Dans cet exemple, les totaux des colonnes sont de 100 %. Comme on l'a expliqué précédemment, ces totaux pourraient figurer dans les rangées. Dans l'exemple du Viet Nam, les totaux de 100 % de chaque rangée indiqueraient comment chaque type de service de santé serait réparti à travers les zones urbaines et rurales du Viet Nam. La troisième possibilité serait que chaque « case » de ce tableau indique la fréquence (en pourcentage) de la probabilité d'une visite à un établissement de soins par une personne d'une zone géographique donnée (urbaine ou rurale); dans ce cas, la somme des pourcentages de toutes les rangées et de toutes les colonnes serait de 100 %. Toutefois, cette option est rarement utilisée, car les distributions conditionnelles sont généralement plus intéressantes. Quoi qu'il en soit, il est utile de fournir assez de données au lecteur pour qu'il puisse en tirer les trois types de fréquences des données présentées sur le tableau.

25. Il y a plusieurs façon d'utiliser les graphiques pour afficher les informations sur la relation entre deux variables qui ne peuvent prendre qu'un faible nombre de valeurs. Lorsque l'on affiche les pourcentages en colonnes ou en rangées, un moyen commode est de présenter plusieurs colonnes verticales dont la somme est égale à 100 %. Chaque colonne représente une valeur particulière de l'une des variables, et la fréquence de distribution de l'autre variable est illustrée sous la forme de zones de différentes couleurs dans chaque colonne. La figure XVI.4 offre un exemple de cette représentation pour l'utilisation des différents types de soins de santé au Viet Nam. Di-

vers logiciels de tableur présentent de nombreuses autres variantes possibles de cette représentation.

Tableau XVI.6

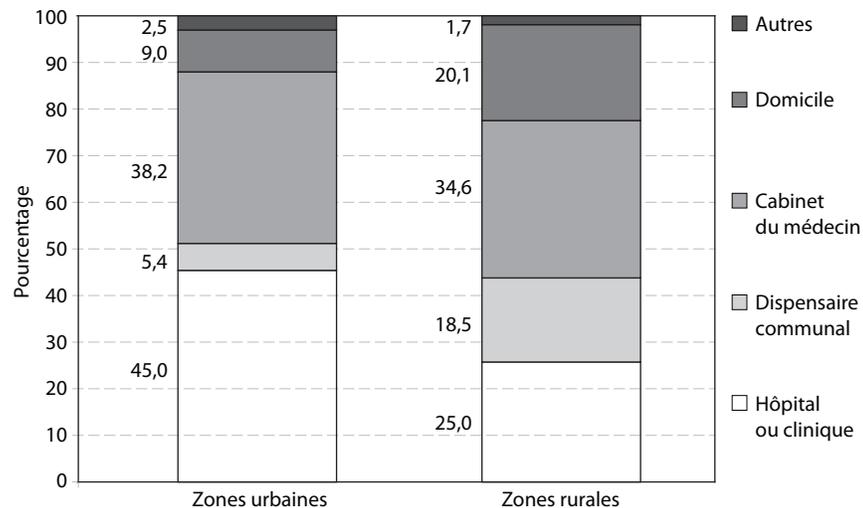
Utilisation des services de santé par les personnes (de tous âges) qui se sont rendues dans un établissement de soins au cours des quatre dernières semaines, par zones urbaines et rurales du Viet Nam, en 1992-1993

Lieu de consultation	Zones urbaines		Zones rurales	
	Fréquence	Pourcentage (erreur type)	Fréquence	Pourcentage (erreur type)
Hôpital ou clinique	251	45,0 (2,1)	430	25,0 (1,0)
Dispensaire communal	30	5,4 (1,0)	318	18,5 (0,9)
Cabinet d'un médecin	213	38,2 (2,1)	595	34,6 (1,1)
Domicile	50	9,0 (1,2)	376	20,1 (1,0)
Autres	14	2,5 (0,7)	29	1,7 (0,3)
Total	558	100,0	1718	100,0

Source : Enquête de 1992-1993 sur les niveaux de vie au Viet Nam.

Figure XVI.4

Utilisation des services de santé par les personnes (de tous âges) qui se sont rendues dans un établissement de soins au cours des quatre dernières semaines, par zones urbaines et rurales du Viet Nam, en 1992-1993 (en pourcentage)



Source : Enquête de 1992-1993 sur les niveaux de vie au Viet Nam.

Note : Taille de l'échantillon : 2 276 ménages.

26. Une variable comptant un petit nombre de catégories/valeurs et une variable numérique pouvant avoir de nombreuses valeurs. Une autre situation à deux variables se présente fréquemment. L'une de ces variables peut compter un petit nombre de catégories ou de valeurs (éventuellement après regroupement pour réduire ce nombre) et l'autre est une variable numérique aux nombreuses valeurs possibles. Ici, la façon la plus courante d'afficher les données est de prendre la moyenne de la variable numérique, en regard de chaque valeur de la variable qui compte un petit nombre de catégories ou de valeurs. On peut ajouter d'autres informations, telles que la valeur médiane et l'écart type. Le tableau XVI.7 en offre un exemple, qui indique le niveau moyen des

dépenses totales des ménages au Viet Nam en 1992-1993 en divisant les ménages entre les sept régions du pays. Ces données pourraient être illustrées par un graphique en colonnes, où chaque colonne (sur l'axe des x) représente une région et la longueur des colonnes (axe des y) est proportionnelle au revenu moyen de chaque région.

Tableau XVI.7

Dépenses totales des ménages par région au Viet Nam, 1992-1993*(en milliers de dong\$ par an)*

Région	Moyenne des dépenses totales (erreurs types entre parenthèses)
Hauts Plateaux du Nord	4 792 (95,5)
Delta du fleuve Rouge	5 306 (110,4)
Centre nord	4 708 (107,7)
Région côtière du Centre	7 280 (234,8)
Hauts Plateaux du Centre	6 173 (373,7)
Sud-Est	10 786 (398,5)
Delta du Mékong	7 801 (167,4)
Ensemble du Viet Nam	6 531 (77,6)

Source : Enquête de 1992-1993 sur les niveaux de vie au Viet Nam.

Note : Taille de l'échantillon : 4 799 ménages.

27. On pourrait également transformer la variable continue en une variable distincte en divisant son spectre en un petit nombre de catégories. Par exemple, il est parfois commode de diviser les ménages par tranches de 20 % selon leur niveau de revenus ou de dépenses. Après cela, on peut utiliser les mêmes méthodes pour présenter les données pour deux variables distinctes, selon la description donnée ci-dessus. On pourrait notamment modifier la figure XVI.4 de manière à montrer cinq colonnes, une pour chaque quintile.

28. *Deux variables numériques pouvant prendre de nombreuses valeurs.* Souvent, les statisticiens offrent des informations sommaires sur deux variables numériques et leur coefficient de corrélation (la covariance des deux variables divisée par la racine carrée du produit des variances). Toutefois, le grand public est généralement peu apte à lire de telles statistiques. On pourrait également afficher les données sous forme graphique avec un point pour chaque observation. Cela montrerait notamment le degré de corrélation du revenu des ménages entre deux périodes et permettrait d'effectuer des observations sur les mêmes ménages dans deux études distinctes (une pour chaque période).

29. Le problème qui se pose avec les graphiques à points épars est que, lorsque l'échantillon est vaste, le graphique comporte trop de points pour permettre une interprétation aisée. On peut éviter cet inconvénient en prélevant un sous-échantillon aléatoire des observations (par exemple, un dixième de celles-ci) pour éviter que le graphique devienne trop confus. L'autre problème qui se pose avec les points épars est de savoir comment les ajuster pour tenir compte des pondérations de l'échantillonnage. La solution simple est de créer des doubles des observations, la pondération de l'échantillonnage étant le nombre de doubles de chaque observation, mais cela ne peut manquer de surcharger le graphique. Par conséquent, après avoir créé les doubles, il ne faut porter qu'un sous-échantillon des observations sur le graphique à points épars.

5. Présentation de statistiques descriptives pour trois variables ou plus

30. En principe, on peut présenter les relations entre trois variables ou plus en utilisant des tableaux et des graphiques. Toutefois, il vaut mieux limiter ces présentations à des cas exceptionnels car cela ajoute des dimensions supplémentaires qui compliquent à la fois la compréhension des relations et les moyens de les présenter sur des tableaux et graphiques simples. Dans la pratique, il est parfois possible de montrer des relations descriptives entre trois variables mais il est presque toujours impossible de montrer de telles relations entre quatre variables ou plus.

31. Pour trois variables, la méthode la plus directe est de désigner une variable comme la variable « conditionnante ». Ou bien cette variable n'a qu'un petit nombre de valeurs distinctes, ou bien, si elle est continue, on doit la « discrétiser » en calculant sa distribution sur un petit nombre d'intervalles disposés sur l'ensemble de son spectre. Lorsque cela est fait, on peut établir des tableaux ou graphiques séparés pour chaque catégorie ou valeur de cette variable conditionnante. Par exemple, supposons que l'on cherche à montrer la relation entre trois variables : le niveau d'instruction du chef de famille, le niveau de revenu du ménage et l'incidence de la malnutrition infantile. On peut le faire en créant un tableau ou un graphique séparé de la relation entre le revenu et un indicateur de l'état nutritionnel des enfants (tel que l'incidence de l'arrêt de croissance) pour chaque niveau d'instruction. Cela peut montrer, par exemple, que la relation entre le revenu et la nutrition de l'enfant est moins défavorable pour les ménages dont le chef a plus d'instruction.

C. Conseils d'ordre général pour la présentation de statistiques descriptives

1. Préparation des données

32. Avant que des chiffres à mettre sur des tableaux et des graphiques soient établis, il faut préparer les données pour l'analyse. Cela implique trois tâches différentes : vérifier les données afin d'éliminer les observations qui pourraient être très inexactes; produire des variables (dérivées) complexes; et documenter avec soin la préparation de l'ensemble de données « officielles » à utiliser pour toute analyse. Durant l'exécution de ces trois tâches, l'attention particulière portée dès le début aux détails peut épargner beaucoup de temps et de ressources par la suite. Dans cette sous-section, nous présenterons un bref aperçu de ces tâches; pour plus de détail, le lecteur est invité à se reporter au chapitre XV.

33. Pratiquement chaque enquête sur les ménages, avec quelque soin qu'elle soit préparée et exécutée, ne peut manquer de donner certaines valeurs aux variables qui ne paraissent pas crédibles. Ces anomalies peuvent aller de la non-réponse ponctuelle (voir chapitre XI) et d'autres erreurs manifestes — par exemple, un enfant de trois ans désigné comme le chef de famille — à des cas beaucoup moins clairs, tels qu'un ménage à très haut revenu mais à niveau de dépenses moyen. Dans beaucoup de cas, les erreurs sont dues à des inexactitudes dans l'entrée des données inscrites sur les questionnaires sur papier, et la vérification devrait commencer par ces questionnaires. Il est facile de remédier à ces erreurs d'entrée des données. Si les données bizarres figurent elles-mêmes sur le questionnaire, il y a plusieurs possibilités. Premièrement, on peut remplacer la valeur bizarre par la mention « manquante ». S'il n'y a qu'un petit nombre de tels cas, on peut exclure ces observations lorsque l'on établit tout tableau ou graphique utilisant cette variable⁴. Si ces cas sont nombreux, les valeurs « manquantes » peuvent être définies comme catégorie distincte d'une variable catégorique,

⁴ Cette option présente l'inconvénient que la taille de l'échantillon est légèrement différente pour chaque tableau. Bien que cela puisse être source de confusion, une note au bas de chaque tableau expliquant que certaines observations ont été éliminées devrait suffire à clarifier la situation.

intitulée « non indiquée ». Deuxièmement, si la plupart des cas sont concentrés sur un petit nombre de ménages, ceux-ci devraient être éliminés de l'enquête. Troisièmement, s'il y a beaucoup d'observations discutables pour de nombreux ménages sur certaines variables, il pourrait être préférable de ne pas présenter les résultats concernant ces variables.

34. Un moyen de remédier aux cas de données manquantes peut être « d'imputer » des valeurs à ces données en appliquant diverses méthodes. Les méthodes d'imputation attribuent des valeurs aux cas inconnus ou non rapportés, ainsi qu'aux cas aux valeurs non plausibles. Certaines de ces méthodes permettent d'imputer une valeur proche de celles recueillies lorsqu'une réponse fait défaut. L'idée est très simple : des ménages ou des personnes qui présentent certaines caractéristiques similaires sont probablement similaires à d'autres titres. Par exemple, les maisons d'un village donné ont des chances d'avoir des murs et des toits semblables à ceux des maisons d'autres zones rurales plutôt qu'à ceux des maisons urbaines. De même, la plupart des membres d'une famille ont la même religion et la même appartenance ethnique. L'équipe chargée de l'enquête doit décider des règles à suivre face aux conditions démographiques, sociales et économiques et aux conditions de logement du pays.

35. Si les méthodes d'imputation sont très utiles, elles peuvent aussi être source de graves problèmes. Les membres de l'équipe chargés de l'analyse des données doivent décider ou non de remplacer les données manquantes en procédant cas par cas ou utiliser un certain type de méthode d'imputation. Il leur faut tenir compte des effets de leur décision sur les tableaux finals. L'imputation de 1 ou 2 % des cas ne devrait pas avoir d'effets sensibles sur les résultats. Si quelque 5 % des cas sont manquants ou en contradiction avec les autres, l'imputation devrait probablement encore pouvoir être envisagée. En revanche, la nécessité d'imputer une proportion beaucoup plus grande de valeurs, disons 1 % ou plus, pourrait très bien rendre la variable impropre à la présentation et à l'analyse, et les résultats relatifs à cette variable ne devraient pas être présentés. Les lecteurs sont invités à se reporter aux chapitres VIII et XI et à leurs références pour plus de conseils sur l'imputation et le traitement à réserver aux valeurs manquantes.

36. Le calcul de variables (dérivées) complexes est un autre aspect de la préparation des données. Dans beaucoup d'enquêtes sur les ménages, on calcule le revenu total ou les dépenses totales des ménages ou les deux sur la base des valeurs d'un grand nombre de variables. Par exemple, on calcule généralement les dépenses totales en ajoutant les dépenses consacrées à l'achat d'au moins 100 articles alimentaires ou autres. Si, en principe, le calcul de ces variables est simple, en réalité, il peut se heurter à un grand nombre de problèmes. Par exemple, quand on calcule les revenus et les dépenses des familles rurales, il peut arriver que les profits agricoles soient négatifs. Lorsque certaines familles affichent des résultats bizarres, il peut être utile d'examiner chacun des éléments qui entrent dans ce calcul. Le problème peut être dû à un ou deux de ces éléments. Toujours en ce qui concerne les profits agricoles, il se peut que le prix d'achat d'un intrant soit exceptionnellement élevé. Dans ce cas, on peut recalculer le profit en utilisant un prix moyen.

37. Malheureusement, la préparation des ensembles de données lorsque se posent des problèmes relève plus de l'art que de la science. On doit prendre des décisions lorsque les choix ne sont pas clairs. Enfin, il importe d'expliquer les choix et, d'une façon plus générale, l'ensemble du processus de transformation des « données brutes » en tableaux et graphiques. Cette explication doit comprendre une brève description du processus et de tous les programmes informatiques utilisés pour manipuler et transformer les données.

2. Présentation des résultats

38. La meilleure façon de présenter les résultats de statistiques élémentaires varie selon le type d'enquête et la clientèle. Néanmoins, on peut donner quelques conseils d'ordre général applicables dans la plupart des cas.

39. Le conseil le plus important est de présenter clairement les résultats. Pour cela, plusieurs recommandations précises s'imposent. Premièrement, toutes les variables doivent être définies clairement et avec précision. Par exemple, lorsque l'on présente des tableaux et des graphiques sur le « revenu » des ménages, la variable « revenu » doit désigner soit le « revenu par personne » soit le « revenu total du ménage », mais jamais le « revenu » sans autre précision. Les variables complexes telles que celles qui indiquent le revenu et les dépenses doivent être définies clairement dans le texte et des notes au bas des tableaux et graphiques. S'agit-il du revenu avant ou après impôt ? Comprend-il la valeur d'un logement s'il est occupé par son propriétaire ? S'agit-il du revenu hebdomadaire, mensuel ou annuel ? Cela doit être indiqué très clairement. Pour beaucoup de variables, il est très utile de présenter dans le texte le libellé du questionnaire d'où provient la variable. Par exemple, pour les données sur l'alphabétisation des adultes, la variable doit être définie très clairement. Elle peut être définie par le nombre d'années de scolarité de la personne, par l'aptitude de celle-ci à écrire son nom ou par son affirmation de son aptitude à lire le journal ou elle peut être fondée sur une sorte de test auquel le répondant est soumis. Les résultats peuvent être différents selon les définitions données.

40. Une autre recommandation est que les distributions en pourcentage de variables distinctes doivent indiquer clairement s'il s'agit de pourcentages de ménages ou de personnes (c'est-à-dire de la population). Souvent, cela donne des résultats différents. Dans beaucoup de pays, les personnes ayant un meilleur niveau d'instruction appartiennent à de petites familles. Cela veut dire que la proportion de la population vivant dans des familles dont le chef a un bon niveau d'instruction est plus faible que celle de ces familles. La troisième recommandation concernant la clarté est que les graphiques devraient indiquer les nombres auxquels correspondent les formes du graphique. Par exemple, le graphique à colonnes de la figure XVI.1 indique les pourcentages pour chacune des trois sources d'éclairage des familles vietnamiennes, et tel est également le cas du diagramme circulaire de la figure XVI.2.

41. Enfin, on peut donner plusieurs autres conseils d'ordre divers. Premièrement, les rapports ne doivent pas présenter trop de tableaux et ceux-ci ne doivent pas contenir trop de chiffres. Parfois, les offices de statistiques présentent des centaines de tableaux donnant quantité de détails infimes qui ont peu de chances d'intéresser la majorité du public, et cela vaut souvent aussi pour les détails que contient un tableau donné. Les personnes qui rédigent le rapport doivent s'interroger sur l'utilité des divers tableaux établis, et s'ils jugent peu utile de présenter un tableau donné ou les informations détaillées qu'il contient, ils devraient éliminer ce tableau ou ces informations. Deuxièmement, il convient de donner les estimations des erreurs d'échantillonnage observées sur certaines des variables les plus importantes recueillies dans le cadre de l'enquête; par ailleurs, il est très utile d'indiquer les intervalles de confiance pour les variables ou indicateurs essentiels. C'est là quelque chose d'évident, mais dont on oublie souvent de tenir compte. Il est important d'indiquer au lecteur le degré de précision des informations découlant de l'enquête sur les ménages. Troisièmement, la taille de l'échantillon doit être précisée pour chaque tableau.

3. Qu'est-ce qui constitue un bon tableau ?

42. La présente sous-section donne des conseils particuliers pour la préparation des tableaux contenant les informations provenant d'une enquête sur les ménages. Lors de l'établissement des tableaux et graphiques, il convient de se conformer au principe général suivant : les informations figurant sur les tableaux devraient être suffisantes pour permettre à l'utilisateur de les interpréter correctement sans avoir à consulter le texte du rapport. Cela est très important car de nombreux utilisateurs des rapports photocopient les tableaux et les utilisent par la suite sans se référer au texte du rapport.

43. Les conseils donnés ci-dessous sont d'ordre général. Quel que soit le cas, l'équipe chargée de l'enquête doit décider des conventions à adopter. Une fois cela fait, elle doit s'en tenir strictement à ces conventions. Toutefois, dans certains cas, elle peut avoir à s'en écarter pour illustrer certains points précis ou certains types particuliers d'analyses statistiques. Enfin, presque toutes les directives données ici pour les tableaux s'appliquent aussi aux graphiques.

44. Le tableau XVI.6 contient les diverses parties d'un bon tableau. Chaque tableau doit avoir un titre clair, des indications géographiques (le cas échéant), des têtes de colonne, des titres pour les rangées, la source des données et toutes notes pertinentes.

45. *Titre.* Le titre doit donner une description succincte du tableau. Cette description doit comprendre : *a)* le numéro du tableau; *b)* la population ou le milieu étudié (y compris l'unité analysée, telle que les ménages ou les personnes); *c)* une indication de ce qui figure dans les rangées; *d)* une indication de ce qui figure dans les colonnes; *e)* le pays ou la région couvert par l'enquête; et *f)* l'année (les années) de l'enquête.

46. En ce qui concerne le numéro du tableau, la plupart des rapports statistiques numérotent leurs tableaux par ordre de présentation, en commençant par le tableau I.1 et en continuant ainsi jusqu'au dernier tableau. Parfois, les pays utilisent des lettres et des chiffres pour différents ensembles de tableaux; par exemple, les tableaux statistiques sur les logements seront numérotés L01, L02, etc. et les tableaux sur la population P01, P02, etc. Si cette procédure est simple et directe, elle présente l'inconvénient de rendre difficiles les additions et les suppressions dans les rapports.

47. Le milieu étudié est la population ou le logement couvert par le tableau. Si celui-ci porte sur toute la population, on peut omettre d'indiquer le milieu dans le titre : il est sous-entendu qu'il s'agit de la population totale. En revanche, si le tableau porte sur une sous-catégorie de la population, par exemple, sur les personnes appartenant à la population active, et que la population active potentielle est définie comme les personnes âgées de 10 ans ou plus, le titre doit alors indiquer « Personnes âgées de 10 ans ou plus ».

48. Le titre du tableau XVI.6 indique également ce qui figure dans les rangées et dans les colonnes du tableau. En particulier, il indique que le tableau présente des informations sur les types de services de santé utilisés (rangées) et présente séparément les informations pour les zones urbaines et rurales (colonnes). La mention du pays ou de la région dans le titre rend immédiatement apparent le milieu géographique considéré. Cela est particulièrement important pour les chercheurs qui veulent comparer les résultats de plusieurs pays. Bien entendu, l'office de statistiques du pays qui recueille les données sait de quel pays il s'agit, mais les personnes qui consultent le tableau dans d'autres pays peuvent avoir besoin de cette information pour savoir de quel pays il s'agit.

49. Enfin, il importe d'indiquer l'année (les années) de l'enquête dans le titre pour rendre la période considérée immédiatement apparente. Parfois, l'office national de statistiques du pays peut décider de présenter les données de deux enquêtes ou plus sur le même tableau. Dans ce cas, il faut indiquer les deux dates; par exemple, « 1990 et 2000 » ou « 1980 à 2000 »; mais une fois cette décision prise, le pays doit s'y tenir.

50. *Identification géographique.* Chaque fois que le même tableau est répété pour des niveaux géographiques inférieurs, chaque tableau doit préciser à quelle région géographique il s'applique. Par exemple, si le tableau XVI.6 devait se répéter pour chacune des sept régions du Viet Nam, le nom de la région devrait figurer entre parenthèses immédiatement en dessous du titre du tableau. On pourrait aussi utiliser des indicateurs non géographiques, par exemple, les appartenances ethniques ou les nationalités couvertes par le tableau.

51. *Têtes de colonne.* Chaque colonne doit avoir une « tête ». Les têtes de colonne peuvent avoir plusieurs « niveaux »; par exemple, au tableau XVI.6, la tête des deux premières colonnes est intitulée « Zones urbaines » et celle des deux dernières colonnes, « Zones rurales »; et aussi bien pour les zones urbaines que rurales, il y a deux sous-titres pour la fréquence des observations et le pourcentage de distribution de ces observations. On peut aussi indiquer les « totaux » ou « sommes » comme pour la première colonne du tableau XVI.3. L'équipe chargée de l'enquête doit choisir une convention concernant l'emplacement de ces colonnes. Traditionnellement, le total vient en dernier, et tous les attributs sont présentés en premier dans la colonne. Toutefois, si un tableau s'étale sur plusieurs pages et compte de nombreuses colonnes d'informations, l'équipe chargée de l'enquête peut préférer indiquer le total en premier (à gauche) pour la série de colonnes. Lorsque le total apparaît en premier, l'utilisateur sait immédiatement quel est le total pour cette série de colonnes, sans avoir à consulter toutes les pages du tableau.

52. Les têtes de colonne et les colonnes de données correspondantes doivent être espacées de manière à limiter le plus possible l'espace laissé en blanc sur la page. Pour déterminer la largeur d'une colonne, il faut tenir compte du nombre de chiffres composant le chiffre maximum et du nombre de lettres composant le nom des attributs figurant dans la colonne, ainsi que du nombre total d'espaces autorisés par la police de caractères utilisée. La police de caractères est très importante et devrait être choisie dès le début du processus de tabulation.

53. *Titres figurant dans les rangées.* L'équipe chargée de l'enquête doit aussi décider des conventions à adopter pour les titres devant figurer dans les rangées. Ces « titres » devraient être justifiés à gauche, et une seule variable devrait figurer sur chaque ligne. Les titres des rangées doivent porter le nom des variables figurant dans la rangée. Ces titres peuvent comprendre des sous-catégories. Par exemple, une catégorie peut être donnée sur deux rangées distinctes, une pour les hommes et une pour les femmes. Il faut établir certaines conventions pour distinguer les différents groupes; la convention est généralement d'indiquer en retrait les différents « niveaux » d'une variable.

54. *Précision des chiffres.* De nombreux tableaux souffrent de la présentation de trop de chiffres significatifs. Lorsqu'ils indiquent des pourcentages, il est presque toujours suffisant de n'inclure qu'un chiffre après la virgule; le fait d'inclure deux décimales ou plus offre rarement des informations utiles et présente trois inconvénients : il distrait le lecteur, gaspille de l'espace et donne une fausse impression de précision. Les nombres à quatre chiffres ou plus ont rarement besoin d'une décimale. Lorsque des nombres élevés sont présentés, ils devraient figurer en « milliers » ou en « millions », afin d'éviter de présenter des nombres avec plus de quatre ou cinq chiffres.

55. *Source.* La source des données devrait être présentée sous la forme du nom complet de l'enquête d'où elles proviennent, généralement au bas du tableau (comme au tableau XVI.6). Toutefois, dans certains cas, les tableaux sont tirés de plus d'une enquête sur un même pays, ou d'enquêtes sur plusieurs pays. Lorsque tel est le cas, les informations figurant dans les sources prennent plus d'importance. La date devrait accompagner le nom de l'enquête. Si la source est un rapport publié, il importe de distinguer la date de publication du rapport de l'année de la collecte des données. Par exemple, un pays peut avoir recueilli des données en 1990 et les avoir publiées en 1992. Dans ce cas, la source devrait indiquer « Enquête de 1990 sur la fécondité, 1992 », où 1992 serait la date de publication.

56. *Notes.* Les notes fournissent les informations immédiates avec lesquelles interpréter correctement les résultats figurant sur le tableau. Par exemple, les notes des tableaux XVI.1 et XVI.2 indiquent que la population sélectionnée comprend toutes les personnes appartenant à un ménage ou vivant en communauté. En plus des notes figurant au bas des tableaux, une série de définitions et d'explications peuvent être données dans le texte accompagnant les tableaux. Ce texte pourrait donner des précisions sur certaines informations; par exemple, il pourrait indiquer comme lieu de naissance l'adresse de la mère avant qu'elle aille à l'hôpital, plutôt que l'hôpital lui-même. Il pourrait également expliquer comment les données ont été recueillies ou doivent être utilisées. Par exemple, si la date de naissance et l'âge sont tous deux indiqués, mais que la date de naissance annule l'âge en cas d'incompatibilité, cette précision doit être donnée pour aider certains utilisateurs, tels que les démographes, à déterminer la meilleure façon d'interpréter les données.

4. Utilisation de pondérations

57. La présente sous-section donne un bref aperçu de l'utilisation de pondération pour l'établissement de tableaux et de graphiques utilisant les données d'une enquête sur les ménages. Pour un traitement beaucoup plus détaillé de cette question, voir les chapitres II, VI, XIX, XX et XXI et les références citées dans ces chapitres.

58. Du point de vue de la pondération, le type le plus simple d'échantillon d'une enquête sur les ménages est l'échantillon « autopondéré ». Dans ce cas, on n'a pas besoin d'utiliser de pondérations dans l'analyse car chaque ménage de la population a la même probabilité de sélection pour l'échantillon. Tel est le cas de l'Enquête de 1992-1993 sur les niveaux de vie au Viet Nam utilisée dans plusieurs des exemples de ce chapitre. Cependant, pour corriger les différences de taux de réponses entre divers types de ménages, il convient de calculer des pondérations. De plus, la plupart des enquêtes sur les ménages ne sont pas autopondérées car elles tirent des échantillons démesurément grands pour certaines parties de la population qu'elles cherchent à étudier de plus près. Pour ces enquêtes, il faut utiliser des pondérations pour tenir compte des différences de probabilité de sélection, de manière à calculer des estimations non faussées des caractéristiques à étudier.

59. Pour être précises, les pondérations doivent comprendre trois éléments. Le premier constitue les « pondérations de base » ou « pondérations de conception ». Ces pondérations tiennent compte de la différence entre les probabilités de sélection de différents groupes de ménages (en l'occurrence, lorsque l'échantillon n'est pas autopondéré), stipulée lors de la conception de l'échantillon initial de l'enquête. Le deuxième élément est l'ajustement pour tenir compte des différences de taux de non-réponse. En effet, dans beaucoup de pays en développement, les ménages plus aisés sont moins enclins à se laisser interroger que les ménages à revenu intermédiaire ou à faible revenu. Les pondérations de base doivent être « gonflées » de l'inverse du taux de réponse de

tous les groupes de ménages. Enfin, dans certains cas, il peut y avoir des « ajustements après stratification ». La logique de la post-stratification est qu'une source de données indépendante, telle qu'un recensement, peut donner des estimations plus précises de la distribution de la population par âge, sexe ou appartenance ethnique. Si les estimations données par l'enquête de ces distributions ne correspondent pas étroitement à celles de la source indépendante, les données de l'enquête peuvent avoir à être re-pondérées pour faire en sorte que les deux distributions coïncident. Pour plus de détails sur les deuxième et troisième éléments, voir Lundström et Särndal (1999).

D. Préparation d'un rapport général (résumé) sur une enquête sur les ménages

60. Les résultats de la plupart des enquêtes sur les ménages sont d'abord diffusés par la publication d'un rapport général contenant quelques détails sur l'ensemble des informations recueillies dans le cadre de l'enquête. Généralement, ces rapports sont diffusés beaucoup plus largement que les rapports plus spécialisés qui utilisent pleinement certains aspects des données. Ces rapports généraux sont parfois appelés « résumés statistiques ». La présente section contient quelques recommandations précises pour la préparation de ces rapports, fondées sur Grosh et Muñoz (1996).

1. Contenu

61. L'élément principal de tout rapport statistique général est un grand nombre de tableaux et graphiques. Ceux-ci doivent contenir tous les principaux types d'informations recueillies dans le cadre de l'enquête; l'analyse approfondie de certaines questions plus spécifiques devrait être laissée pour des rapports spéciaux plus focalisés. Les tableaux doivent s'accompagner de quelques explications, juste assez pour préciser le type d'informations figurant sur ces tableaux. S'il n'est pas nécessaire que ce rapport tire des conclusions particulières de politique générale, certaines interprétations peuvent être suggérées comme sujets de recherche pour l'avenir.

62. Les informations les plus élémentaires peuvent être décomposées par région géographique, par sexe, voire par âge. Si l'enquête contient des données sur les revenus et les dépenses, celles-ci peuvent aussi être fractionnées par groupes de revenus ou de dépenses. Dans certains pays, il y a de grandes différences entre ces différents groupes, et la nature de ces différences peut être étudiée plus avant dans des tableaux supplémentaires. Dans d'autres pays, certaines de ces différences ne sont pas grandes et n'ont pas besoin d'une présentation plus détaillée.

63. Outre les résultats de l'enquête sur les ménages, le rapport général devrait consacrer plusieurs pages à décrire l'enquête elle-même et indiquer notamment la taille et la conception de l'échantillon, la date de démarrage et la date d'achèvement de l'enquête et fournir certains détails sur la façon dont les données ont été recueillies. Le questionnaire ou les questionnaires ayant servi à l'enquête devrait/devraient figurer en annexe au rapport.

2. Processus

64. Un bon rapport statistique général est réalisé par une équipe dont plusieurs membres devraient de préférence avoir l'expérience de rapports précédents. Certains s'intéressent aux aspects techniques de la production de tableaux et de graphiques, tandis que d'autres s'occupent principalement du contenu et du texte d'accompagnement. Les techniciens peuvent choisir le logiciel de statistiques qu'ils connaissent le

mieux, car la plupart de ces logiciels peuvent produire les chiffres nécessaires à l'établissement des tableaux et graphiques. Toutefois, l'estimation des erreurs types exige généralement un logiciel conçu spécialement à cette fin, car les échantillons des enquêtes sur les ménages sont pratiquement toujours trop complexes pour pouvoir être traités correctement par les logiciels de statistiques courants (voir l'analyse de cette question au chapitre XXI).

65. Les membres de l'équipe chargée du contenu du rapport doivent se réunir avec des experts des organismes publics pour examiner les sujets à inclure dans le rapport. Ainsi, les tableaux et graphiques ont des chances de présenter les données les plus utiles à ces organismes. Il pourrait être bon également de consulter les institutions internationales d'aide, qui pourraient elles aussi juger les données utiles à la planification de leurs programmes (voir l'examen plus général de la façon de former une bonne équipe au chapitre III).

E. Conclusions

66. Ce chapitre a fourni une introduction à la présentation de statistiques descriptives simples utilisant les données d'enquêtes sur les ménages. L'examen a été très général et s'est limité à un niveau très élémentaire. Si tout ce qui a été présenté n'est que le fruit d'un peu de bon sens, il est recommandé aux analystes d'user eux aussi de leur bon sens lorsqu'ils se trouvent en présence de questions particulières concernant l'analyse de leurs enquêtes. On peut utiliser des méthodes plus complexes, dont certaines sont examinées dans les chapitres suivants, pour analyser les données d'une enquête sur les ménages. En fin de compte, l'analyse des données d'une enquête sur les ménages doit être fonction du thème principal et des objectifs de l'enquête, et il incombe aux chercheurs de consulter des ouvrages et journaux spécialisés pour y trouver des conseils propres aux thèmes étudiés.

RÉFÉRENCES

- Frankenberg, Elizabeth (2000). Community and price data. Dans *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 Years of the Living Standards Measurement Study*, M. Grosh et P. Glewwe, eds. New York, Oxford University Press, pour la Banque mondiale.
- Grosh, Margaret, et Juan Muñoz (1996). *A Manual for Planning and Implementing the Living Standards Measurement Study Survey*. Living Standards Measurement Study Working Paper, No. 126. Washington, D.C., Banque mondiale.
- Lundström, S., et C. E. Särndal (1999). Calibration as a standard method for the treatment of non-response in sample surveys, *Journal of Official Statistics*, vol. 13, n° 2, pp. 305-327.
- Tufte, Edward (1983). *The Visual Display of Quantitative Information*. Cheshire, Connecticut, Graphics Press.
- Wild, C. J., et G. A. F. Seber (2000). *Chance Encounters: A First Course in Data Analysis and Inference*. New York, Wiley.

Chapitre XVII

Utilisation d'enquêtes sur les ménages à thèmes multiples pour améliorer les politiques de lutte contre la pauvreté dans les pays en développement

PAUL GLEWWE

Department of Applied Economics
 Université du Minnesota
 Saint Paul, Minnesota, États-Unis d'Amérique

RÉSUMÉ

Le présent chapitre montre comment les enquêtes sur les ménages peuvent être utilisées par les chercheurs et les pouvoirs publics des pays en développement pour formuler des politiques de lutte contre la pauvreté. Il commence par des analyses descriptives relativement simples, soulignant la contribution essentielle des données des enquêtes sur les ménages : ces enquêtes fournissent les informations qui permettent de déterminer qui est pauvre et comment se caractérisent les pauvres. Puis il examine des analyses multivariates plus complexes fondées sur des techniques à base de régressions multiples. Pour chaque type d'analyse, il présente des exemples de la façon dont les données des enquêtes sur les ménages peuvent servir à formuler des politiques de lutte contre la pauvreté.

Termes clés : pauvreté, formulation de politiques, analyses descriptives, analyses multivariates.

A. Introduction

1. La plupart des pays en développement reconnaissent que l'un des principaux objectifs du développement économique et social est de faire reculer et, à terme, d'éliminer la pauvreté. Si tous les gouvernements ont le même but, les politiques qu'ils mettent en œuvre pour combattre la pauvreté ne sont pas nécessairement les mêmes. La nature même de la pauvreté et ce qui la caractérise varient d'un pays à un autre et il est donc normal que les politiques soient, elles aussi, différentes les unes des autres.

2. Le présent chapitre est beaucoup trop bref pour donner un examen approfondi des nombreuses façons dont les politiques gouvernementales peuvent agir sur la pauvreté dans les pays en développement. [Voir Lipton et Ravallion (1995) et la Banque mondiale (2001) pour une analyse détaillée de cette question.] Néanmoins, on peut donner un aperçu général de ces politiques, que l'on divisera, aux fins du présent chapitre, en quatre grandes catégories. La première regroupe les politiques macroéconomiques, c'est-à-dire les politiques à l'échelle de l'économie qui ont des répercus-

sions sur la croissance économique et la stabilité. Les politiques macroéconomiques les plus importantes ont trait au niveau général de l'imposition et des dépenses publiques, aux politiques monétaires (qui influent sur les taux d'intérêt et le taux d'inflation), aux politiques économiques internationales (qui conditionnent le taux de change et les flux de commerce et de capitaux extérieurs) et aux politiques à l'égard des banques et autres institutions financières. La deuxième catégorie de politiques gouvernementales comprend celles qui influent sur les prix, telles que les taxes et subventions sur certains biens et services donnés. La mise en place par l'État de services publics et d'infrastructure, notamment de cliniques, d'écoles et de réseaux de transport et de communication, représente la troisième catégorie générale de politiques gouvernementales. La dernière a trait aux programmes gouvernementaux conçus pour apporter une aide directe aux pauvres. Comme exemples de ces politiques, on peut citer le Programa de Educación, Salud y Alimentación (PROGRESA) au Mexique, qui distribue des allocations en espèces aux familles pauvres dont les enfants fréquentent assidûment l'école, et le programme de distributions de bons d'alimentation en Jamaïque, qui distribue des bons pouvant servir à l'achat de produits alimentaires dans les magasins locaux. Ces quatre types de politiques peuvent avoir des effets sensibles sur la pauvreté.

3. L'impact sur la pauvreté de l'une quelconque de ces politiques dépend des caractéristiques et du comportement des pauvres et, dans certains cas, de ceux de la population non pauvre. En effet, l'incidence sur la pauvreté des subventions gouvernementales qui permettent d'abaisser les prix de certains produits alimentaires dépend de la mesure dans laquelle les pauvres achètent ces produits. Par conséquent, le gouvernement a besoin d'informations sur les caractéristiques et le comportement des pauvres du pays pour déterminer les politiques qui peuvent aider le plus efficacement à combattre la pauvreté. Les enquêtes sur les ménages fournissent des renseignements cruciaux à cet égard.

4. Presque tous les pays en développement, même les plus pauvres, effectuent certains types d'enquêtes sur les ménages, notamment sur les revenus et les dépenses et sur la main-d'œuvre, ainsi que des enquêtes démographiques et sanitaires. Ces enquêtes leur fournissent une mine d'informations qui leur permettent de mieux comprendre la nature de la pauvreté et les effets probables des politiques gouvernementales sur les pauvres. Ce chapitre précise comment les pays en développement peuvent utiliser les enquêtes sur les ménages pour formuler leurs politiques de lutte contre la pauvreté. La section B commence par montrer ce que l'on peut tirer de simples statistiques descriptives calculées à partir des données d'une enquête. La section C, qui examine certaines méthodes plus complexes fondées sur l'analyse multivariée, est suivie d'une brève conclusion.

B. Analyse descriptive

5. Pour assurer l'efficacité des politiques et programmes gouvernementaux destinés à aider les pauvres, il faut savoir si ces politiques et programmes profitent vraiment aux pauvres et quels sont leurs effets. Malheureusement, il est rare que les pays en développement disposent de telles informations. En effet, certaines mesures qui tendent à favoriser la croissance économique peuvent contribuer à accroître à des degrés divers les revenus des différentes professions. Dès lors se pose la question de savoir quelles professions sont les plus courantes chez les pauvres. On peut faire la même remarque au sujet des politiques de prix. Par exemple, l'incidence sur les pauvres de la décision d'un gouvernement de relever les taxes sur le pétrole dépend de la quantité de produits pétroliers consommés par les familles pauvres. La situation est la même en ce qui concerne la question de savoir s'il faut construire de nouvelles écoles ou cliniques

dans certaines régions du pays : ces régions ont-elles une concentration relativement forte de familles pauvres ? Enfin, les responsables de la gestion de tout programme de prestations en faveur des pauvres, qu'il s'agisse de services ou de transferts en nature ou en espèces, ont besoin de savoir quelle est la proportion des bénéficiaires du programme qui sont vraiment pauvres et quelle est la proportion des pauvres du pays qui en bénéficient.

6. Malheureusement, nombre de pays en développement sont peu informés de la localisation et des caractéristiques des pauvres et n'ont donc qu'une très faible notion de la mesure dans laquelle leurs politiques et programmes gouvernementaux profitent aux pauvres ou leur nuisent. Les enquêtes sur les ménages peuvent combler beaucoup de ces déficits d'informations. La présente section examine comment cela peut se faire, en utilisant de nombreux exemples tirés de pays en développement. Bien que beaucoup de données de ces enquêtes pour comprendre la pauvreté soient très faciles à utiliser et se limitent à l'établissement de simples tableaux et graphiques, ces informations sont souvent beaucoup plus utiles que celles qui proviennent d'analyses complexes.

1. Définir la pauvreté

7. Avant d'analyser l'impact des politiques gouvernementales sur les pauvres, il faut bien préciser ce que l'on entend par pauvre, autrement dit, il faut définir la pauvreté. Les avis ne sont pas toujours unanimes sur cette définition. Cependant, on s'accorde généralement à reconnaître qu'il existe en principe un niveau minimum « acceptable » auquel les individus et les ménages sont en droit d'aspérer pour mener une vie décente. La majeure partie du débat sur la pauvreté est axée sur les besoins matériels, plutôt que sur les libertés politiques, les droits de l'homme et le bien-être psychologique, et nous ferons de même dans ce chapitre. Les besoins matériels les plus évidents, et qui font donc l'objet d'un large consensus, sont : *a*) une alimentation adéquate; *b*) un logement décent; et *c*) l'accès à l'eau potable et à des moyens sanitaires d'évacuation des déchets. La plupart des observateurs ajouteraient également à cela des possibilités d'éducation de base et des soins préventifs simples. D'autres plaideront pour un ensemble plus complet de biens et de services, ajoutant par exemple des activités culturelles et récréatives, mais les avis sont partagés sur ce qu'il convient d'inclure dans cet ensemble, ou même sur le point de savoir si ces types de biens et de services devraient être tout simplement inclus.

8. Les philosophes, les économistes et autres spécialistes des sciences sociales peuvent consacrer, et ont en fait souvent consacré, beaucoup de temps à débattre ce que devrait comprendre l'ensemble minimum de biens et de services auxquels un individu ou une famille devrait avoir accès pour ne pas être considéré comme pauvre. Une fois qu'un tel ensemble est accepté, l'absence de consommation de l'un de ses éléments particuliers peut être utilisée comme « indicateur » de pauvreté. L'approche plus pratique, adoptée par de nombreux économistes, tend à souligner que presque tous les articles qui composent cet ensemble coûtent de l'argent et que la vraie question n'est pas de savoir la composition exacte de cet ensemble mais son coût monétaire. Selon cette approche, on fixe un « seuil de pauvreté » exprimé sous la forme d'un montant donné et l'on définit comme pauvre tout ménage dont le revenu ou les dépenses se situent en dessous de ce montant. En fait, le point de départ de toute définition du seuil de pauvreté utilisée dans les pays en développement est un ensemble de biens et de services qui correspondent aux besoins minimaux. Par exemple, cet ensemble pourrait comprendre un élément composé des produits alimentaires correspondant au minimum de besoins nutritionnels et aux profils nationaux de consommation d'aliments. L'opé-

ration suivante serait de calculer le coût de cet ensemble. Dans le reste de ce chapitre, nous supposerons que telle est l'approche suivie; pour un examen plus approfondi de la façon de déterminer le seuil de pauvreté, voir Ravallion (1998).

2. Construction d'un profil de pauvreté

9. Une fois que l'on a établi une définition utilisable de la pauvreté, exprimée en fonction du revenu ou des dépenses des ménages, on peut décrire les pauvres en utilisant les données de l'enquête sur les ménages. Cette description est souvent appelée « profil de pauvreté » et se fonde sur les données de l'enquête sur les ménages relatives aux revenus et/ou aux dépenses pour calculer le pouvoir d'achat total de chaque ménage (le total de ses revenus ou de ses dépenses). Les pauvres sont définis comme les ménages dont le pouvoir d'achat se situe en dessous du seuil de pauvreté.

10. On peut tirer implicitement une leçon et une question du paragraphe précédent. La leçon est que l'analyse de la pauvreté a besoin que les données de l'enquête sur les ménages fournissent des informations suffisamment précises sur le total des revenus et/ou des dépenses des ménages. Sans de telles données, cette analyse est difficile car il faut trouver un autre moyen de classer les ménages comme pauvres ou non pauvres. S'il est vrai que l'enquête peut fournir certaines informations utiles en dehors même de ces données, on peut apprendre bien davantage des enquêtes sur les ménages qui recueillent des données sur les revenus et/ou les dépenses. La question qui se pose est la suivante : si l'enquête fournit des données à la fois sur les revenus et les dépenses, lesquelles des deux doit-on utiliser ? En général, on préfère les données sur les dépenses car elles sont souvent plus précises que les données sur les revenus et parce qu'elles sont, théoriquement, plus étroitement liées au bien-être du ménage, car le revenu est parfois utilisé pour rembourser des dettes ou pour épargner pour l'avenir et, en tant que tel, ne reflète pas nécessairement la situation de bien-être actuelle.

11. Lorsque l'on construit un profil de pauvreté, la première tâche est de déterminer qui sont les pauvres. Sans les données des enquêtes sur les ménages, les décideurs et autres observateurs n'ont souvent qu'une idée très limitée de ce que sont les pauvres et de ce qui les caractérise. De surcroît, certaines de leurs perceptions peuvent être erronées. Par exemple, nombre de décideurs et autres observateurs passent la majeure partie de leur temps dans de grandes agglomérations urbaines et ne se font une idée des pauvres que d'après ce qu'ils voient en ville, alors que, dans presque tous les pays, l'incidence de la pauvreté est beaucoup plus grande dans les zones rurales. Autrement dit, lorsque l'on utilise les données d'une enquête sur les ménages, la première tâche doit être d'estimer l'incidence de la pauvreté, de décrire le cadre dans lequel vivent les pauvres, défini en zones urbaines ou rurales et par région, et de calculer certaines des caractéristiques fondamentales des pauvres. Il importe de vérifier les taux de pauvreté par groupe ethnique et religieux, selon le niveau d'instruction et selon la profession. Il est également utile d'examiner les conditions de logement des pauvres, ainsi que la notion de propriété des biens de production. Avec ces informations et d'autres encore, on peut commencer à donner des conseils utiles aux décideurs.

12. Un rapport récent de la Banque mondiale (1999) sur la pauvreté au Viet Nam, qui a estimé que, en 1998, 37 % de la population de ce pays était pauvre, offre un exemple des traits essentiels qui caractérisent les pauvres. Au Viet Nam, 79 % des pauvres travaillent dans l'agriculture, et presque tous à titre indépendant. Un autre fait essentiel est que la pauvreté est beaucoup plus répandue parmi les groupes ethniques minoritaires : alors que ces groupes ne représentent que 14 % de la population vietnamienne, ils représentent 29 % des pauvres du pays.

13. L'endroit où ils vivent est l'un des principaux aspects qui caractérisent les pauvres. Les décideurs aimeraient connaître de préférence l'incidence de la pauvreté dans chaque grande ville, agglomération et district rural. Malheureusement, la taille de l'échantillon d'une enquête typique sur les ménages est généralement de l'ordre de 3 000 à 15 000 ménages, ce qui est trop peu pour permettre des estimations précises de la pauvreté dans chacun de ces contextes. Toutefois, si l'on dispose aussi des données d'un recensement récent, on peut combiner ces données à celles d'une enquête sur les ménages pour produire des estimations de la pauvreté pour des zones géographiques beaucoup plus petites. Il s'agit essentiellement d'estimer la relation entre diverses variables « prédictives » et le revenu ou les dépenses des ménages, en utilisant les données de cette enquête. Les variables prédictives utilisées figurent également dans le recensement. Le fait qu'avec une estimation de la relation de prédiction on peut utiliser les données du recensement pour simuler la distribution des dépenses dans des zones géographiques relativement petites permet d'estimer l'incidence de la pauvreté dans ces zones. Hentschel *et al.* (1998) en offre un exemple en utilisant des données de l'Équateur. Pour un examen plus approfondi des méthodes employées, voir Rao (2002) et Kalton (2002).

14. Le dernier point important en ce qui concerne la définition de la pauvreté et la construction de profils de pauvreté est que l'on veut souvent comparer la pauvreté à différentes périodes dans le même pays ou à une même période dans différents pays. Pour ce faire, il importe que les données de l'enquête sur les ménages utilisées pour définir les dépenses ou le revenu soient recueillies de la même façon dans le temps ou à travers les pays. De très légères différences dans la conception du questionnaire ou d'autres changements dans la méthode de collecte des données peuvent souvent entraîner des modifications importantes mais totalement fallacieuses, et souvent imprévues, des estimations. À vrai dire, il peut être impossible de faire de telles comparaisons si les données recueillies ou la façon dont elles sont analysées, ou les deux, ne sont pas les mêmes dans les enquêtes comparées. Autrement dit, tous changements dans la façon dont les données sont recueillies pour les variables qui définissent la pauvreté doivent être considérés avec beaucoup de précaution, si l'on veut limiter le risque que les changements observés ne soient dus qu'aux procédures statistiques et non pas à de véritables changements. C'est pourquoi il est généralement préférable de modifier le moins possible la façon dont les données sont recueillies.

3. Utilisation des profils de pauvreté pour une simple analyse des politiques

15. Le fait de savoir où vivent les pauvres et de connaître certaines de leurs caractéristiques essentielles est la première condition à remplir pour donner des conseils aux décideurs. Bien entendu, les programmes destinés spécifiquement à aider les pauvres doivent être mis en œuvre là où se trouvent les plus fortes concentrations de pauvres, mais on peut faire beaucoup plus sur le plan des programmes en analysant convenablement les statistiques concernant ces populations. La présente sous-section décrit quatre types d'informations de base sur les pauvres, qui peuvent servir à tirer les leçons de l'impact de diverses politiques les concernant.

16. *Comment les pauvres gagnent leurs revenus.* Comme on l'a expliqué précédemment, l'une des façons dont les politiques gouvernementales agissent sur les pauvres est en agissant sur leurs revenus. Dès lors, la question importante qui se pose est de savoir ce que font les pauvres pour gagner leur vie. La première question à poser est : les pauvres travaillent-ils à leur compte ou pour un employeur ? Dans beaucoup de pays, la grande majorité des pauvres travaillent à leur compte comme agriculteurs,

artisans ou commerçants. Par définition, les pauvres qui travaillent à leur compte ne sont pas directement touchés par les mesures concernant les salariés, telles que les modifications des lois sur le salaire minimum ou la mise en œuvre d'un régime de « sécurité sociale » ou d'assurance maladie qui ne s'applique qu'aux salariés.

17. Comme beaucoup de pauvres sont de petits exploitants agricoles, la question importante est de savoir ce qu'ils produisent, et quelle quantité de leur production est destinée à la vente. Dans un exemple tiré de la Côte d'Ivoire, Glewwe et De Tray (1990) ont constaté que beaucoup d'agriculteurs ivoiriens sont producteurs de coton, alors que la production de coton est très rare parmi les agriculteurs non pauvres. Ainsi, les décisions gouvernementales touchant le prix du coton influent au premier chef sur les pauvres du pays.

18. *Profils de consommation des pauvres.* La situation économique des pauvres est aussi déterminée par les prix des biens et des services qu'ils consomment. Par exemple, au Ghana, moins de 1 % des 20 % les plus pauvres de la population possède une moto ou une voiture (Glewwe et Twum-Baah, 1991). Cela implique un très faible effet direct d'une hausse du prix de l'essence sur ces éléments de la population ghanéenne, mais un effet indirect dû à la hausse des prix des transports publics.

19. D'une façon plus générale, les données sur la consommation d'articles alimentaires et autres et sur l'accès à l'électricité et l'eau courante représentent une mine d'informations pour les décideurs. Lorsqu'une taxe ou une subvention est envisagée sur un type de bien particulier, il faut examiner ces données afin de voir quels seraient ses effets sur les pauvres. Les politiques de taux de change influent aussi sur les prix, de sorte qu'il convient également de tenir compte de la consommation par les pauvres de biens importés. L'exemple donné ci-dessus du Ghana illustre bien ce point, dans la mesure où tous les produits pétroliers consommés par le pays sont importés.

20. *Services utilisés par les pauvres.* Les subventions à la santé et à l'éducation sont souvent justifiées, du moins en partie, par les avantages qu'elles procurent aux pauvres. Toutefois, il y a beaucoup de types de services de santé et de niveaux d'éducation. Les données qui indiquent qui utilise ces services permettent de vérifier l'état de pauvreté des bénéficiaires de telle ou telle mesure.

21. Le Viet Nam en offre un exemple récent. Gertler et Litvack (1998) ont constaté que l'élément type des 20 % les plus pauvres de la population faisant en moyenne une consultation externe par an dans un hôpital de la santé publique et deux dans un dispensaire communal. En revanche, un élément type des 20 % les plus riches faisait en moyenne quatre ou cinq consultations externes par an dans un hôpital de la santé publique et seulement une dans un dispensaire communal. La principale raison de cette disparité est que les hôpitaux de la santé publique se trouvent principalement dans les zones urbaines, alors que 90 % des pauvres du pays vivent dans les zones rurales. La conséquence évidente de ces chiffres est que les subventions aux dispensaires communaux profitent davantage aux pauvres qu'aux non-pauvres, tandis que les subventions aux hôpitaux profitent beaucoup plus aux non-pauvres qu'aux pauvres.

22. *Participation aux programmes.* Enfin, les données des enquêtes sur les ménages ont une autre utilité immédiate : elles permettent de déterminer qui participe aux divers programmes de l'État destinés à aider les pauvres. Pour ce faire, il faut qu'elles posent au moins une question précise sur la participation des ménages à ces programmes, et qu'elles fournissent des données sur les revenus et les dépenses qui permettent de classer les ménages en pauvres et non-pauvres. Ces données, rares par le passé, sont aujourd'hui de plus en plus courantes dans les enquêtes, car les concepteurs ont pris conscience de leur valeur.

23. La Jamaïque (Grosh, 1991) offre un exemple de l'utilisation d'une enquête sur les ménages pour évaluer le ciblage d'un programme. Les données de cette enquête ont montré que les bons d'achat de produits alimentaires avaient beaucoup plus de chances, ce qui n'est peut-être pas étonnant, d'être utilisés par les familles pauvres que par les autres. Curieusement, les avantages de subventions alimentaires générales tendent à profiter principalement aux familles plus aisées. Cette information a été présentée au gouvernement à la fin des années 1980, et au début des années 1990, les avantages du programme de bons d'achat de produits alimentaires ont été doublés tandis que les subventions alimentaires ont été abandonnées.

24. La dernière remarque générale sur l'analyse descriptive simple est que presque toutes les enquêtes sur les ménages sont fondées sur des échantillons complexes plutôt que sur des échantillons aléatoires. Par conséquent, les sous-groupes de population présentant un intérêt particulier, tels que les pauvres, sont suréchantillonnés, ce qui implique qu'il faut utiliser des pondérations d'échantillonnage pour produire des estimations objectives des statistiques descriptives de base. De plus, le calcul des erreurs types doit tenir compte de la conception de l'échantillon. Ces points étant traités de façon plus détaillée au chapitre XVI et dans d'autres chapitres de cet ouvrage, le lecteur est invité à consulter ces chapitre avant d'entreprendre une analyse descriptive.

C. Analyse de régression multiple des données d'enquêtes sur les ménages

25. Les exemples donnés ci-dessus de l'utilisation des enquêtes sur les ménages sont fondés sur des statistiques très simples qui peuvent être calculées par quiconque peut utiliser un logiciel de statistique simple. Toutefois, les leçons qui peuvent en être tirées sont parfois trop simplistes en ce sens qu'elles ne tiennent pas compte des réactions de comportement à ces politiques. Par exemple, si une taxe sur une denrée agricole donnée est supprimée parce que cette denrée est produite couramment par les pauvres, les ménages non pauvres risquent eux aussi de commencer à produire cette denrée à mesure que son prix augmente, de sorte que certains des avantages de cette mesure peuvent aller aux familles non pauvres¹. De même, une taxe sur une denrée particulière peut avoir un effet très négatif sur les pauvres s'ils consomment une grande quantité de cette denrée, mais si une autre denrée similaire n'est pas taxée, les pauvres peuvent être tentés simplement de la substituer à l'autre au prix d'une légère diminution de leur bien-être. L'éducation offre un autre exemple. Le fait que les enfants pauvres d'un pays donné fréquentent rarement le lycée tend à indiquer qu'il y a peu d'avantages pour les pauvres à réduire les frais de scolarité de ces écoles; or, il est possible qu'une telle réduction favorise un fort accroissement de la présence d'enfants pauvres dans ces établissements. Cette possibilité implique donc qu'un simple examen des chiffres de fréquentation des lycées tendrait à sous-estimer l'avantage pour les pauvres d'une telle réduction.

26. Les enquêtes sur les ménages peuvent être utilisées pour estimer comment le comportement des ménages change lorsque les politiques changent. Cela n'est pas une tâche facile, car elle nécessite des types d'analyse beaucoup plus complexes. Les méthodes les plus couramment utilisées pour effectuer une telle estimation sont celles de l'analyse de régressions multiples. Les méthodes les plus complexes utilisent souvent les données d'enquêtes sur les ménages spécialement conçues pour recueillir les données précises nécessaires à cette analyse. Cela est nécessaire car ces méthodes nécessitent souvent des données que ne produisent pas les enquêtes typiques sur les ménages. La présente section décrit trois façons courantes d'utiliser les données d'enquêtes sur les ménages pour estimer comment certaines politiques influent sur le

¹ Lorsque la taxe est en place, le prix perçu par le producteur est inférieur au prix payé par le consommateur, la différence étant le montant de la taxe. Lorsque celle-ci est supprimée, le prix au producteur est égal au prix au consommateur, et dans presque tous les cas, cela veut dire que le prix perçu par le producteur augmente et que le prix payé par le consommateur diminue.

comportement des ménages. Deaton (1997) offre un examen plus approfondi de cette question.

1. Analyse de la demande

27. Les économistes estiment souvent l'impact des prix et des revenus des ménages sur les achats de biens et de services. Cette recherche est appelée analyse de la demande. L'idée générale est que pour tout bien (i), les achats de ce bien (q_i) par un ménage sont fonction du revenu (y) du ménage, du prix (p_i) de ce bien et des prix de tous les autres biens, ce qui s'exprime par l'équation :

$$q_i = f(y, p_1, p_2, \dots, p_i, \dots, p_n) + \varepsilon \approx \beta_0 + \beta_1 p_1 + \beta_2 p_2 + \dots + \beta_i p_i + \dots + \beta_n p_n + \beta_{n+1} y + \varepsilon.$$

La fonction $f(y, p_1, p_2, \dots, p_i, \dots, p_n)$ est une représentation très générale de la façon dont le revenu et les prix influent sur la demande des ménages, où ε reflète l'impact d'autres facteurs causaux (et peut-être une variation aléatoire de q_i qui n'a rien à voir avec aucun facteur causal). On simplifie couramment l'analyse de la demande en supposant une représentation linéaire, représentée ici par le terme situé à droite du symbole « \approx », qui indique que cette simplification n'est qu'une approximation. Si ε est sans corrélation avec y et les variables de prix, on peut alors utiliser les moindres carrés ordinaires (OLS) pour obtenir des estimations objectives des coefficients (les β) du revenu (y) et des prix (p_i) dans cette relation linéaire. Dans la réalité, cette supposition risque de n'être pas valable, et de nombreuses autres questions d'estimation risquent de compliquer l'analyse. Pour plus de renseignements sur l'estimation du système de la demande, la référence classique est Deaton et Muellbauer (1980). Cette question a été étudiée plus récemment par Pollack et Wales (1992) et Lewbel (1997).

28. Pour se rendre compte de la façon dont l'analyse de la demande fournit plus d'informations que les simples statistiques descriptives, considérons l'impact d'une taxe sur une denrée alimentaire importée, telle que le blé. (Souvent, les économies en développement taxent les produits importés car ces taxes sont relativement faciles à administrer; et du fait de leur climat tropical, beaucoup de pays en développement ne peuvent avoir de blé que par l'importation.) Supposons que le prix actuel du kilo de farine de blé est de 10 et que les ménages pauvres typiques en consomment 60 kilos par an. Si l'on suppose que le prix à l'importation est fixé au cours international, une taxe de 50 % sur les importations de blé porterait le prix de la farine de blé à 15, ce qui veut dire que le ménage pauvre typique paierait 300 (5×60) de taxes additionnelles. Bien entendu, cette analyse, fondée sur de simples statistiques descriptives, suppose que les ménages pauvres continueraient d'acheter la même quantité de farine de blé, une fois la taxe imposée. En fait, il est probable que ces ménages réduiraient leur consommation de farine de blé et augmenteraient leur consommation d'autres produits de consommation courante (tels que le riz, le maïs ou le manioc). Les estimations de l'analyse de la demande permettent de calculer l'ampleur de ce changement de comportement. Supposons que l'équation du paragraphe précédent décrive la demande de blé, de sorte que q_i représente le nombre de kilos de farine de blé achetés chaque année par chaque ménage et p_i représente le prix du kilo de farine de blé. Si $\beta_i = -3$, une hausse de 5 de ce prix entraînerait une diminution de 15 de la consommation de farine de blé, de sorte que la consommation annuelle d'un ménage pauvre serait en moyenne de 45 kilos. Cela veut dire que le ménage pauvre paierait en moyenne 225 (5×45) de taxes additionnelles au lieu de 300. Si cet exemple est très simple, il souligne la nécessité de tenir compte de la réaction des ménages lorsque l'on analyse l'impact de certaines mesures.

29. Deaton, Parikh et Subramanian (1994) offrent un exemple de l'utilisation de l'analyse de la demande pour étudier l'impact des mesures gouvernementales sur

les pauvres. Les auteurs estiment un système d'équations de la demande pour plus de 10 produits alimentaires. Ils calculent l'impact global des hausses des prix des produits alimentaires sur le bien-être social du pays, ainsi que l'impact de ces changements sur la condition des pauvres. L'un des résultats intéressants de leur analyse est que les hausses du prix du riz ont des effets moins négatifs sur le bien-être des pauvres que les hausses des prix des céréales secondaires, dont les pauvres sont plus gros consommateurs. Par conséquent, les taxes sur le riz ont moins d'effets nuisibles sur les pauvres que celles qui frappent les céréales secondaires.

2. Utilisation des services sociaux

30. Les programmes de santé et d'éducation peuvent offrir de nombreux avantages aux ménages pauvres, mais la participation ne signifie pas nécessairement qu'ils en tirent des avantages substantiels. Certains de ces programmes peuvent être inefficaces. Dans le domaine de la santé, les décideurs veulent savoir si la participation se traduit effectivement par un meilleur état de santé. Dans le domaine de l'éducation, ils aimeraient savoir combien les enfants apprennent en fréquentant l'école. De nombreuses études ont été effectuées à l'aide des données d'enquêtes sur les ménages dans les pays en développement pour mesurer le succès des programmes de santé et d'éducation dans la réalisation de leurs objectifs.

31. Une analyse de l'impact de certaines caractéristiques des écoles sur l'acquisition de savoir des élèves et, par là même, sur leurs salaires à venir, offre un exemple récent de l'utilisation de régressions multiples. Glewwe (1999) a étudié cette question en estimant l'impact des caractéristiques des écoles et des ménages sur les résultats scolaires des enfants, mesurés à leurs notes, à partir des données d'une enquête sur les ménages au Ghana. L'équation utilisée était du type suivant :

$$T_i = \beta_0 + \beta_1 MED_i + \beta_2 FED_i + \beta_3 y_i + \beta_4 IQ_i + \beta_5 SC_{1i} + \beta_6 SC_{2i} + \dots + \varepsilon,$$

où T_i est la note obtenue par l'élève i , MED_i et FED_i sont respectivement les niveaux d'instruction de la mère et du père de l'élève i , y_i est le revenu du ménage auquel appartient l'élève i et les variables SC représentent un grand nombre de caractéristiques de l'école et des enseignants. Une telle équation est très difficile à calculer (voir Glewwe, 2002); mais une fois les valeurs de β estimées, ces valeurs indiquent comment différentes caractéristiques concernant l'école et les enseignants influent sur les résultats de l'élève. En comparant ces impacts avec les coûts liés aux diverses caractéristiques de l'école et des enseignants, on peut se faire une idée des types de dépenses d'éducation qui sont les plus efficaces.

32. L'analyse effectuée par Glewwe des données de l'enquête sur les ménages au Ghana a permis de constater que la réparation des toits des salles de classe qui avaient des fuites et l'installation de tableaux dans les classes qui n'en avaient pas amélioreraient sensiblement les résultats des élèves et de l'école (ceux-ci mesurés au nombre d'années de scolarité des élèves). De simples calculs des taux de rentabilité financière de tels « investissements » dans la qualité des écoles indiquaient des taux très élevés, parfois de 25 % ou plus.

3. Impact de certains programmes gouvernementaux

33. S'il est facile d'utiliser les données d'une enquête sur les ménages pour savoir si un ménage ou un individu donné participe à un certain type de programme conçu pour aider les pauvres, il est plus difficile de déterminer dans quelle mesure cette participation améliore effectivement le bien-être des intéressés. En effet, le problème tient au fait que la participation peut avoir d'autres effets qui diminuent ce bien-

être. Par exemple, un programme « aliments contre travail » peut offrir de l'emploi à des pauvres, mais les avantages provenant de l'accroissement du revenu s'accompagnent d'un surcroît de travail, avec ce que cela peut comporter pour la santé de ces individus. De même, lorsque les ménages reçoivent des bons d'achat d'aliments pour accroître leur consommation alimentaire, ils n'utilisent pas nécessairement ces bons à cette fin et peuvent très bien en détourner une partie ou la totalité du produit pour d'autres usages. Pour mesurer l'impact de ces programmes sur le comportement des ménages, il faut procéder à une analyse économétrique minutieuse et compliquée pour comprendre tous les effets de la participation au programme et, en fin de compte, l'impact général de cette participation sur le bien-être des ménages.

34. Un exemple récent nous est fourni par une étude de Jacoby (2002) qui a examiné l'impact de programmes de repas scolaires aux Philippines. Cette étude cherchait à déterminer si la fourniture de repas aux enfants des écoles ne se traduisait pas par un rationnement alimentaire de ces enfants chez eux par leurs parents. Si la plupart des économistes s'attendaient à un tel rationnement, Jacoby n'en a trouvé aucune manifestation. En effet, il a constaté que la participation au programme de repas scolaires ne changeait rien à la consommation alimentaire de l'enfant chez lui, ce qui veut dire que la consommation alimentaire totale de ces enfants augmentait en fait de la quantité représentée par ces repas scolaires.

D. Résumé et conclusions

35. Les enquêtes sur les ménages représentent une source précieuse d'informations à utiliser par les décideurs et les concepteurs de programmes pour déterminer si certaines politiques et certains programmes profitent aux ménages pauvres. Pour être utile, une enquête doit recueillir des données sur les revenus ou les dépenses pour permettre de classer les ménages comme pauvres et non pauvres, et des données indiquant comment une politique ou un programme particulier peut influencer sur la condition des ménages. Récemment encore, les enquêtes sur les ménages étaient souvent conçues pour d'autres usages. Toutefois, dans les années 1980 et 1990, de nouvelles enquêtes sur les ménages ont été conçues spécialement pour fournir ce type d'informations. Parmi les principales, on peut citer les enquêtes de la Banque mondiale pour l'étude de la mesure des niveaux de vie. Ces enquêtes sont présentées brièvement dans Grosh et Glewwe (1998) et de façon très approfondie dans Grosh et Glewwe (2000). Néanmoins, même les enquêtes types conçues à d'autres fins pourraient être rendues beaucoup plus utiles pour l'analyse de la pauvreté si on leur ajoutait quelques questions supplémentaires. Par exemple, il serait très utile d'ajouter des questions sur la participation aux programmes nationaux de lutte contre la pauvreté (tels que les programmes d'emploi rural ou de bons d'achat d'aliments) aux enquêtes courantes sur le revenu et les dépenses des ménages.

36. Ce chapitre a donné au lecteur un aperçu de la façon d'utiliser les enquêtes sur les ménages pour concevoir des politiques visant à faire reculer la pauvreté dans les pays en développement. L'analyse est brève, il est vrai, en raison de l'espace limité dont dispose cette publication. Pour un examen plus approfondi des questions traitées, il est recommandé au lecteur de se reporter aux ouvrages et études cités ici.

RÉFÉRENCES

Banque mondiale (1999). *Viet Nam: Attacking Poverty*. Joint Report of the Government of Viet Nam – Donor – NGO Poverty Working Group. Hanoi.

- _____ (2001). *World Development Report 2000/2001: Attacking Poverty*. New York, Oxford University Press.
- Deaton, Angus (1997). *The Analysis of Household Surveys: A Microeconomic Approach to Development Policy*. Baltimore, Maryland, Johns Hopkins University Press.
- _____, et John Muellbauer (1980). *Economics and Consumer Behaviour*. Cambridge, Royaume-Uni et New York, Cambridge University Press.
- _____, Kirit Parikh et Shankar Subramanian (1994). Food demand patterns and pricing policy in Maharashtra: an analysis using household-level survey data. *Sarvekshana*, vol. 17, pp. 11-34.
- Gertler, Paul, et Jennie Litvack (1998). Access to health care during the transition: the role of the private sector in Viet Nam. Dans *Household Welfare and Viet Nam's Transition*, D. Dollar, P. Glewwe et J. Litvack, eds. Washington, D.C., Banque mondiale.
- Glewwe, Paul (1999). *The Economics of School Quality Investments in Developing Countries*. Londres, Macmillan.
- _____ (2002). Schools and skills in developing countries: education policies and socioeconomic outcomes. *Journal of Economic Literature*, vol. 40, n° 2, pp. 436-482.
- _____, et Dennis de Tray (1990). The poor during adjustment: a case study of Côte d'Ivoire. Dans *Macroeconomic Policy Reforms, Poverty, and Nutrition*, P. Pinstrip-Andersen, ed. Ithaca, New York, Cornell Food and Nutrition Policy Program Monograph, n° 3.
- _____, et Kwaku Twum-Baah (1991). *The Distribution of Welfare in Ghana, 1987-88*. Living Standards Measurement Study Working Paper, No. 75. Washington, D.C., Banque mondiale.
- Grosh, Margaret (1991). *The Household Survey as a Tool for Policy Change: Lessons from the Jamaican Survey of Living Conditions*. Living Standards Measurement Study Working Paper, No. 80. Washington, D.C., Banque mondiale.
- _____, et Paul Glewwe (1998). Data Watch: the World Bank's Living Standards Measurement Study Household Surveys. *Journal of Economic Perspectives*, vol. 12, n° 1, pp. 187-196.
- _____ (2000). *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 Years of the Living Standards Development Study*. New York, Oxford University Press (pour la Banque mondiale).
- Hentschel, Jesko, et al. (1998). *Combining Census and Survey Data to Study the Spatial Dimensions of Poverty*. Policy Research Working Paper, No. 1928. Washington, D.C., Banque mondiale.
- Jacoby, Hanan (2002). Is there an intra-household "flypaper effect"? evidence from a school feeding programme. » *Economic Journal*, vol. 112, n° 476 (janvier), pp. 196-221.
- Kalton, Graham (2002). Models in the practice of survey sampling (revisited). *Journal of Official Statistics*, vol. 18, pp. 129-154.
- Lewbel, Arthur (1997). Consumer demand systems and household equivalence scales. Dans *Handbook of Applied Economics*, vol. II, *Microeconomics*. H. Pesaran et P. Schmidt, eds. Oxford, Royaume-Uni, Blackwell.
- Lipton, Michael, et Martin Ravallion (1995). Poverty and policy. Dans *Handbook of Development Economics*, vol. 3. J. Behrman et T.N. Srinivasan, eds. North Holland.
- Pollak, Robert, et Terence Wales (1992). *Demand System Specification and Estimation*. Oxford, Royaume-Uni, Oxford University Press.
- Rao, J. N. K. (2002). *Small Area Estimation*. New York, Wiley.
- Ravallion, Martin (1998). *Poverty Lines in Theory and Practice*. Living Standards Measurement Study Working Paper, No. 133. Washington, D.C., Banque mondiale.

Chapitre XVIII

Méthodes multivariates pour la construction d'indices

SAVITRI ABEYASEKERA

Centre de services statistiques

Université de Reading

Reading, Royaume-Uni de Grande-Bretagne et d'Irlande du Nord

RÉSUMÉ

Par leur nature même, les enquêtes engendrent des structures de données multivariates. Tout en reconnaissant l'intérêt d'approches simples de l'analyse des données d'enquêtes, le présent chapitre illustre les avantages qu'offre une analyse plus approfondie, ayant recours à l'application de techniques multivariates, pour certains sous-groupes de population. Il existe à présent des logiciels qui permettent aux chercheurs d'appliquer ces méthodes plus perfectionnées.

Ce chapitre décrit un ensemble de situations où les méthodes multivariates peuvent jouer un rôle dans la construction d'un indice et aux premiers stades de l'étude de données sur des sous-ensembles particuliers de données d'enquête, avant qu'une analyse plus approfondie ne soit effectuée pour répondre à certains objectifs spécifiques de l'enquête. Il se concentre principalement sur les méthodes impliquant l'étude simultanée de plusieurs variables clés. À cet égard, les méthodes multivariates permettent une étude plus approfondie de tendances éventuelles affichées par les données, facilitent une représentation graphique de relations complexes entre de nombreuses données et offrent des moyens de réduire la dimensionnalité des données pour des analyses sommaires et complémentaires. L'analyse de la construction d'indices utilise une interprétation plus large des méthodes multivariates qui englobent les méthodes de type régression.

Tout au long de ce chapitre, on s'efforce de donner un aperçu général des méthodes multivariates afin de permettre d'apprécier leur valeur pour la construction d'indices d'un point de vue pratique. Ce chapitre s'adresse à la fois aux personnes qui travaillent à de vastes enquêtes sur les ménages et aux chercheurs travaillant à des projets de recherche et de développement qui n'ont peut-être pas une grande expérience de l'application des approches de l'analyse décrites ici. L'utilisation de ces méthodes est illustrée par des exemples appropriés et un examen de la façon dont les résultats peuvent s'interpréter.

Termes clés : construction d'indices, méthodes multivariates, principaux éléments, analyse par grappes.

A. Introduction

1. Lors de l'analyse des données d'une enquête, la plupart des analystes utilisent généralement des approches statistiques directes. Les approches les plus courantes sont les tableaux à une dimension, à deux dimensions ou à dimensions multiples, et l'utilisation de représentations graphiques telles que diagrammes à barres, graphiques à courbes, etc. On trouvera un aperçu de ces approches et une bonne analyse des aspects nécessitant une certaine attention lors du processus d'analyse des données dans Wilson et Stern (2001) et aux chapitres XV et XVI de la présente publication. Toutefois, dans certains cas, on peut avoir besoin d'une analyse allant au-delà des simples résumés. Nous examinerons dans le présent chapitre un de ces types d'analyse.

2. Les méthodes multivariées traitent simultanément plusieurs variables (Krzanowski et Marriott, 1994a et 1994b; Sharma, 1996). Au strict sens statistique du terme, elles portent sur l'étude collective d'un groupe de variables, tenant compte des relations qui existent entre les variables de ce groupe. Toutefois, de nombreux chercheurs utilisent également le terme « multivariate » dans l'application des techniques de régressions multiples car ces techniques font intervenir plusieurs variables explicatives (prédictrices) aux côtés de la principale variable étudiée (par exemple, Ruel, 1999). Une fois encore, l'examen collectif de plusieurs variables offre l'avantage de permettre d'établir des corrélations. Les méthodes de régression, qui portent essentiellement sur la modélisation d'une variable clé, sont examinées de façon plus approfondie au chapitre XIX. Ici, nous nous pencherons principalement sur l'étude conjointe de plusieurs variables de mesure, comme étape préliminaire d'une interprétation plus large des méthodes multivariées dans l'examen de la construction d'indices.

3. Les techniques multivariées sont souvent perçues comme techniques « avancées » nécessitant un haut niveau de connaissance statistique. S'il est vrai que les aspects théoriques de nombreuses procédures multivariées et leur application peuvent être décourageants même pour les statisticiens, ces procédures n'en ont pas moins un rôle utile à jouer dans l'analyse des données des enquêtes effectuées dans les pays en développement. Nous examinerons d'abord l'utilisation effective de ces méthodes : *a*) en tant que moyen exploratoire d'étudier les tendances affichées par les données; *b*) afin d'identifier les groupements naturels de la population pour une analyse plus approfondie; et *c*) afin de réduire la dimensionnalité du nombre de variables en cause. Nous considérons cet examen comme une phase préliminaire devant conduire à la construction d'indices à partir des variables observées au niveau des ménages, par exemple pour créer des indicateurs de pauvreté [voir, par exemple, Sahn et Stifel (2000)].

4. La section B présente un aperçu général des techniques multivariées sous la forme d'une étude collective d'un groupe de variables de sortie. Elle est suivie de cinq sections traitant de l'application d'un certain nombre d'exemples donnés à titre d'illustration. La section finale présente des conclusions sur la valeur et les limites de ces techniques. Nous avons limité le plus possible les détails techniques, et mettons davantage l'accent sur la compréhension des concepts et leur interprétation. Le lecteur qui désire mieux comprendre ces techniques est invité à consulter Everitt et Dunn (2001) et Chatfield et Collins (1980).

B. Quelques restrictions quant à l'utilisation des méthodes multivariées

5. Dans ce chapitre, nous présentons l'utilisation des approches multivariées en tant que procédures descriptives utiles aux premiers stades de l'exploration des

données et de la construction d'indices. Toutefois, avant d'appliquer ces méthodes, il importe de souligner dès le départ qu'une analyse appliquée de l'ensemble complet de données d'une enquête nationale sur les ménages a peu de chances de déboucher sur des conclusions utiles, compte tenu de l'inévitable diversité des ménages de tout pays. On risque de perdre des informations précieuses si l'on combine dans une même analyse les populations urbaines et rurales et les différentes zones agroécologiques, car les moyens de subsistance des ménages de strates aussi différentes peuvent être très variés. Les techniques décrites dans ce chapitre ne doivent donc être utilisées qu'après un examen minutieux de la structure des données, afin que l'on puisse identifier les différents secteurs ou sous-strates de la population auxquels ces méthodes peuvent s'appliquer, sans perdre de vue les principaux objectifs de l'enquête.

6. Même à l'intérieur de chaque sous-strate ou lorsque l'analyse doit porter sur l'échantillon tout entier, il importe de tenir compte des pondérations associées aux unités sélectionnées. Si ces pondérations varient sensiblement pour les données analysées, l'utilisation d'un logiciel qui n'offre pas de moyens de tenir compte des pondérations de l'échantillon peut conduire à des conclusions erronées. En pareils cas, la pondération des unités sélectionnées d'après les pondérations de l'échantillon, par exemple en utilisant la définition de *WEIGHT* donnée dans SAS (2001) ou la commande *aweight* dans STATA (2003) permet de résoudre cette difficulté pour ce qui est des méthodes couvertes aux sections C, D, E et F. De nombreux autres logiciels permettent de tenir compte des pondérations d'échantillonnage pour ce qui est des méthodes décrites à la section G. Lorsque l'on n'utilise pas les pondérations d'échantillonnage, il faut faire preuve de prudence dans l'interprétation des résultats, qui peuvent faire l'objet de certaines distorsions.

C. Aperçu des méthodes multivariates

7. L'idée essentielle sur laquelle repose l'utilisation des méthodes multivariates dans les enquêtes est la simplification; par exemple, la volonté de ramener un ensemble vaste et parfois complexe de données à quelques mesures sommaires significatives ou de cerner les principales caractéristiques et toutes tendances intéressantes qu'affichent les données. Le but recherché est souvent le tâtonnement : ces méthodes peuvent aider à formuler des hypothèses intéressantes pour le chercheur plutôt qu'à les tester. Nombre d'approches utilisent des méthodes qui ne présupposent de distribution statistique sous-jacente pour aucune des variables. Cependant, il faut faire preuve d'une certaine prudence quant aux types de données utilisées (par exemple, intervalle-échelle, comptages, binaire); nous reviendrons sur cette question à l'endroit voulu dans ce chapitre.

8. Le point de départ est une matrice de données où les rangs représentent des cas (les unités de l'échantillon) et les colonnes, les variables. Parfois, les rangs sont plus intéressants; par exemple, s'ils représentent les familles d'agriculteurs, il peut être intéressant de grouper les ménages par catégorie de revenu sur la base d'un certain nombre de critères socio-économiques représentés par certaines colonnes de la matrice de données. Dans d'autres cas, les colonnes peuvent être l'élément primordial; par exemple, lorsqu'un ensemble de données correspondant à un thème particulier doit être combiné pour former un indice composite pour une analyse plus approfondie.

9. Dans les sections ci-dessous, nous nous concentrons sur quatre principales approches du traitement de données multivariates dans les enquêtes effectuées dans les pays en développement. Les trois premières peuvent être considérées comme des techniques exploratoires utilisées pour la construction de l'indice. Tout d'abord, nous examinerons des procédures graphiques et des mesures sommaires propres à nous faire mieux comprendre les données. Ensuite, nous nous pencherons sur deux procédures

multivariates courantes, l'analyse par grappes et l'analyse de la composante principale (ACP), car ces deux procédures ont un rôle important à jouer dans la construction de l'indice. On examinera la seconde plus à fond à la section G, avec d'autres façons de construire l'indice, en prenant l'interprétation plus large des méthodes multivariates utilisée par de nombreux chercheurs. Tout au long de cet examen, nous supposerons qu'un sous-ensemble approprié des données de l'enquête a été choisi pour l'analyse et que le but recherché en soumettant ces données à une procédure multivariante est d'intégrer une étape exploratoire dans une analyse destinée à répondre à un objectif plus large.

10. Bien entendu, il existe de nombreuses autres méthodes multivariates qui pourraient être envisagées dans certaines situations particulières. Le tableau XVIII.1 présente un ensemble de ces méthodes en donnant une brève description de chacune. Ce chapitre se limite aux trois premières, car il cherche à se concentrer sur l'exploration des données comme première étape nécessaire à la construction d'un indice. Ces trois méthodes sont aussi celles qui présentent le plus d'intérêt pour l'analyse des données d'enquêtes. Avec l'application plus large donnée au terme « multivariante » dans notre discussion sur la construction d'indices, elles constituent de précieux outils méthodologiques pour l'analyse de ces données. Les autres méthodes figurant au tableau XVIII.1 peuvent être utiles dans certains cas précis où elles conviennent aux objectifs de l'enquête. Toutefois, elles débordent le cadre de ce chapitre, qui se propose de n'offrir qu'une présentation générale de certaines des méthodes les plus simples.

Tableau XVIII.1
Quelques techniques multivariates et leur but

Technique multivariante	But de la technique
1. Méthodes multivariates descriptives	Étude de données, identification de tendances et de relations.
2. Analyse de la composante principale	Réduction de la dimension par l'établissement de nouvelles variables (la principale composante) comme combinaisons linéaires des variables de l'ensemble multivariante.
3. Analyse par grappes	Identification de groupements naturels parmi les cas ou les variables.
4. Analyse de facteurs	Modélisation de la structure des corrélations entre les variables de l'ensemble de réponses multivariates par rapport à un ensemble de facteurs communs.
5. Analyse multivariante de variance (MANOVA)	Extension de l'analyse univariante de la variance à l'étude simultanée de plusieurs variates. Il s'agit de fractionner la somme totale des carrés et la matrice des produits croisés entre un ensemble de variates selon la conception expérimentale.
6. Analyse des discriminants	Déterminer une fonction qui permette de séparer deux groupes d'individus ou plus.
7. Analyse de corrélation canonique	Étude de la relation entre deux groupes. Elle suppose la formation de paires de combinaisons linéaires des variables de l'ensemble multivariante de façon que chaque paire produise elle-même la plus forte corrélation entre les individus des deux groupes.
8. Mise à échelle multidimensionnelle	Construction d'une « carte » indiquant une relation spatiale entre un certain nombre d'objets, en commençant à partir d'une table des distances entre ces objets.

D. Graphiques et mesures sommaires

11. Dans toute analyse de données, une compréhension préalable des données est un premier pas essentiel. Un examen attentif des données donne une première impression de leur signification et de leurs profils de distribution; il permet d'identifier les éventuelles aberrations (les observations qui s'écartent du profil des autres données); de déceler des tendances; et d'indiquer à l'utilisateur si certaines variables accusent une plus grande variabilité que d'autres [voir, par exemple, Tufte (1983) et Everitt et Dunn (2001)].

12. Comme dans un ensemble d'analyses univariates, certaines mesures sommaires telles que moyennes et écarts types des données et tables de fréquences de données binaires et catégoriques sont souhaitables. Des paires de variables peuvent donc être envisagées pour déterminer les associations entre variables. Au stade préliminaire, il serait raisonnable de considérer les données par « paquets », par exemple, deux paquets, l'un regroupant les données quantitatives (continues ou distinctes) et l'autre, les données qualitatives (catégoriques et binaires). Pour le premier, des points épars (groupés par paires) seraient significatifs, tandis que, pour le second, des tableaux à deux entrées, également appariés, conviendraient, éventuellement combinés à un certain degré d'association et à l'utilisation d'une statistique chi-carrée. Le cas échéant, les points épars peuvent aussi être représentés à l'aide de symboles différents permettant d'indiquer les sous-ensembles de données identifiés par une variable catégorique.

13. La plupart des logiciels de statistiques offrent des possibilités de points matriciels, par exemple la procédure PLOT dans SAS (2001), le menu *Graph/Graphics* dans SPSS pour Windows (SPSS, 2001) et GenStat pour Windows (GenStat, 2002). Il s'agit de représentations graphiques où les points épars de toutes les paires de variables peuvent être représentés ensemble, permettant ainsi de se faire rapidement une idée de la façon dont chaque variable est liée à chaque autre variable dans l'ensemble de données multivariates examiné.

14. À titre d'exemple, la figure XVIII.1 présente une matrice, produite à partir de SPSS (2001), qui illustre les relations entre quatre variables pour 50 villages de l'État de Gujarat, en Inde, selon que ces villages ont ou n'ont pas une coopérative laitière. Ces variables étaient : la population du village, sa superficie et le nombre de têtes de bétail bovin et de buffles, ces variables n'étant que quelques-unes parmi un groupe plus large. Les données proviennent d'une étude de base effectuée avant l'introduction d'un programme de promotion de la formation en santé animale. Les coordonnées de chaque ensemble de points sont déterminées par l'axe parallèle aux cellules diagonales. Par exemple, les trois cases de la première rangée ont toutes la population du village comme axe horizontal et la superficie, le nombre de têtes de bétail bovin et de buffles comme axes verticaux. Les mêmes ensembles de points apparaissent dans la première colonne, mais les axes sont inversés. Les données semblent laisser apparaître une aberration, clairement visible dans la première rangée, qui correspond à un village à très forte population. On remarque une certaine association entre toutes les paires de variables. On voit également que les valeurs élevées de toutes les variables étudiées ont plus de chances de correspondre à des villages ayant une coopérative laitière qu'à des villages n'en ayant pas.

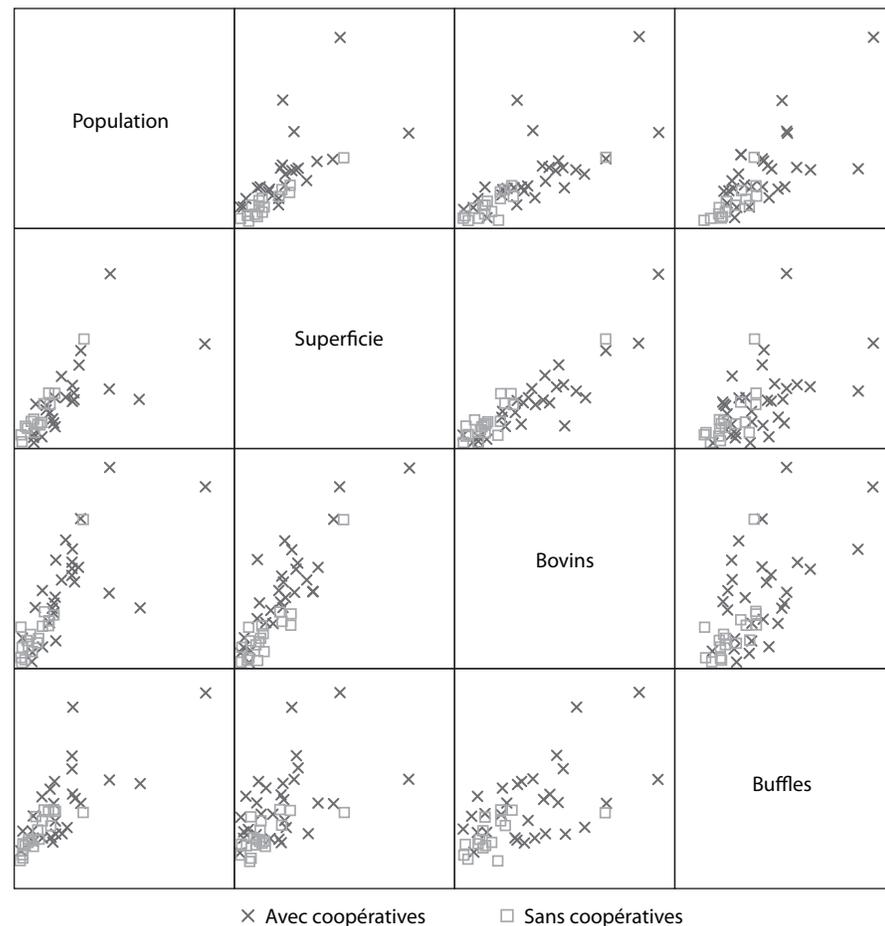
15. Si la matrice laisse apparaître des paires de variables indiquant des tendances intéressantes ou des aberrations, il est recommandé d'effectuer une nouvelle mesure sous la forme d'un repérage à deux entrées de points épars, en tenant compte des pondérations d'échantillonnage liées à chaque point. Les graphiques où chaque point est représenté par une bulle, dont l'étendue est proportionnelle à la pondération de l'échantillon (Korn et Graubard, 1998), sont particulièrement utiles et permettent une interprétation plus significative. Par exemple, une valeur aberrante ayant une forte pondération aura nécessairement un impact plus marqué qu'une autre à pondération plus faible. Il y a plusieurs autres moyens de tenir compte de la conception de l'échantillon lorsque l'on obtient des points épars, notamment en effectuant un sous-échantillonnage des données selon une probabilité proportionnelle aux pondérations de l'échantillon puis en marquant les points sans tenir compte des pondérations ou en

appliquant des méthodes de « kernel smoothing ». Pour plus de détails, se reporter à Korn et Graubard (1998).

16. Il existe de nombreuses autres méthodes de représentation graphique de données multivariées. Par exemple, Manly (1994) montre comment plusieurs objets, décrits à l'aide de plusieurs variables, peuvent être représentés de trois façons différentes pour illustrer le profil des valeurs des variables. L'ouvrage d'Everitt et Dunn (2001) contient un excellent chapitre sur les nombreuses représentations graphiques, notamment les boxplots bivariés, les coplots et les graphiques en treillis, et l'étude de Jongman, Ter Braak et Van Tongeren (1995) donne une démonstration de l'utilisation de biplots. On ne peut donner plus de détails dans cet ouvrage, mais nous invitons le lecteur à consulter les références citées ci-dessus pour plus de renseignements. Il convient toutefois de noter que ces méthodes graphiques sont les plus utiles lorsqu'elles sont utilisées pour certains sous-groupes particuliers de population.

Figure XVIII.1

Exemple de représentation matricielle de quatre variables



E. Analyse par grappes

17. L'analyse par grappes (Everitt, Landau et Leese, 2001) est une technique d'exploration des données qui vise généralement à déterminer les groupements naturels parmi les unités d'échantillonnage (par exemple, les répondants, les exploitations, les ménages), de manière que les unités de chaque groupe (grappe) soient semblables

les unes aux autres et que les unités différentes soient placées dans des groupes différents. Dans certains cas, il peut être utile de mettre en grappe des variables; par exemple, lorsque l'on ne choisit qu'une ou deux variables de chaque grappe, afin que l'analyse se fonde sur un moins grand nombre de variables. L'analyse par grappes est donc un moyen utile d'étudier des données et/ou d'en réduire le nombre. Elle peut aussi aider à formuler des hypothèses et servir dans certaines autres situations données.

Exemple 1

18. À titre d'illustration, considérons une étude visant à évaluer l'efficacité d'un ensemble de stratégies peu coûteuses de lutte contre les ennemis des cultures destinées à être utilisées par les familles d'agriculteurs pauvres d'une certaine région. Supposons qu'une étude de base des agriculteurs susceptibles de participer à de futurs essais sur le terrain soit effectuée afin de déterminer : *a*) le profil socio-économique de ces familles d'agriculteurs; *b*) leurs pratiques actuelles de lutte contre les ennemis des cultures; et *c*) leurs perceptions de l'action de ces ravageurs sur leurs cultures. Nous nous concentrons ici sur le premier de ces trois buts et examinerons comment l'analyse par grappes peut aider à faire un choix judicieux de différents groupes d'agriculteurs pour l'étude principale comportant les essais sur le terrain.

19. Lors de l'enquête de base, on a mesuré un grand nombre de variables socio-économiques, afin de stratifier les familles d'agriculteurs sur la base de ces variables. L'une des approches consiste, par exemple, à choisir deux variables clés et à former des strates définies sur la base de combinaisons de catégories liées à ces deux variables. Par exemple, si les variables choisies sont le sexe du chef de famille (homme/femme) et le niveau de sécurité alimentaire (faible, moyen, élevé), on obtient six strates.

20. L'inconvénient de cette approche est qu'elle ne tient pas compte des caractéristiques socio-économiques des ménages. Une approche multivariate permet de considérer simultanément de nombreuses variables. L'analyse par grappes appliquée aux familles d'agriculteurs sur la base de toutes les variables socio-économiques pertinentes est un moyen plus efficace de stratifier les ménages en un certain nombre de grappes, de telle manière que chaque grappe représente un groupe socio-économique distinct d'agriculteurs. Cela est important dans la mesure où les mêmes recommandations concernant les stratégies de lutte contre les ennemis des cultures ne conviennent pas nécessairement à toutes les familles d'agriculteurs. Une première classification de celles-ci en grappes permet de mieux choisir les différents types d'agriculteurs en vue de leur participation à l'étude d'un ensemble de ces stratégies. Elle permet aussi de se concentrer sur les caractéristiques propres aux différentes grappes, et d'étudier ainsi l'interaction entre ces caractéristiques et les stratégies recommandées. On trouvera une illustration de cette application de l'analyse par grappes dans l'étude d'Orr et Jere (1999).

21. Deux décisions s'imposent avant l'exécution d'une analyse par grappes. Premièrement, il faut déterminer le degré de similarité (ou de divergence) entre les unités mises en grappes. La mesure de la similarité utilise les informations tirées de plusieurs variables pour donner une valeur numérique reflétant le degré de « proximité » entre deux paires d'unités. La mesure de la divergence est le contraire et reflète la distance qui sépare deux paires d'unités données. Lorsque toutes les variables sont quantitatives ou ne comprennent tout au plus que quelques variables catégoriques *ordonnées*, il peut être opportun d'utiliser la matrice de distance euclidienne¹. Toutefois, les données d'enquêtes comprennent souvent des variables catégoriques binaires et non ordonnées. Pour ces types de données, diverses mesures de similarité ont été proposées. Par exemple, pour produire une mesure de similarité entre deux variables

¹ La distance euclidienne peut être considérée comme reflétant simplement la signification normale de la « distance » dans un espace multidimensionnel.

binaires, il faut d'abord établir une tabulation croisée des données du type du tableau 2×2 ci-dessous :

	0	1
0	<i>a</i>	<i>b</i>
1	<i>c</i>	<i>d</i>

22. On peut alors obtenir une mesure de similarité $(a + d)/(a + b + c + d)$, que l'on appelle coefficient de simple concordance. Il y a également le coefficient $d/(b + c + d)$. On trouvera un ensemble d'autres mesures dans Krzanowski et Marriott (1994b). Pour obtenir une mesure satisfaisante de la similarité avec des types de données mixtes, voir Gower (1971). Dans la pratique, si l'on doit utiliser un grand nombre de variables de différents types pour la mise en grappes, il peut être préférable de procéder à un certain nombre d'analyses de grappes différentes, en considérant chaque fois les variables du même type, puis de déterminer si les divers ensembles de grappes obtenus sont similaires. Cela permet d'effectuer une validation croisée de la composition des grappes.

23. Une fois que l'on a déterminé la mesure de la divergence ou de la similarité, il faut décider de la méthode de mise en grappes. Là encore, les logiciels de statistiques offrent de nombreuses options. Par exemple, SPSS (2001) en offre sept (à savoir, liens entre groupes, liens au sein d'un groupe, voisin le plus proche, etc.). Certaines de ces procédures sont dites agglomératives, c'est-à-dire qu'au départ les n unités mises en grappes forment n grappes avec un membre par grappe, puis sont combinées de façon séquentielle selon leur similarité avec les membres d'autres grappes. L'autre option est celle de la méthode divisive, selon laquelle les n unités forment une seule grappe, qui est ensuite divisée d'une manière séquentielle jusqu'à ce que l'on obtienne une solution satisfaisante. Dans un cas comme dans l'autre, il faut veiller à prendre la bonne décision en ce qui concerne la formation des grappes. Everitt, Landau et Leese (2001) donnent une analyse approfondie de ces questions.

Exemple 2

24. On se trouve en présence d'un cas particulier lorsque toutes les variables sont binaires. La procédure peut être relativement simple et consister à utiliser une hiérarchie de grappes. À titre d'illustration, nous n'utiliserons que quelques observations inspirées par une petite enquête qui portait sur 74 agriculteurs inclus dans un programme de recherche au niveau des exploitations. Les données correspondant à un certain nombre de variables enregistrées lors de visites des exploitations sont présentées au tableau XVIII.2 pour huit agriculteurs. Les variables correspondent à des réponses par oui (+) et par non (-). L'un des objectifs était de déterminer si les exploitations pouvaient être groupées en quelques grappes sur la base de ces caractéristiques.

25. Une fois encore, aux fins d'illustration et pour maintenir simples les détails de la construction, considérons la formation d'une matrice de similarité utilisant le nombre de + que tous les ensembles de deux variables ont en commun. Les résultats sont présentés au tableau XVIII.3. On peut ainsi former un ensemble de grappes en considérant les huit exploitations comme constituant huit grappes, puis en fusionnant les grappes les plus proches jusqu'à ce que toutes les exploitations ne forment plus qu'une seule grappe.

26. La figure XVIII.2 donne une représentation graphique de la matrice de similarité de l'exemple ci-dessus. Ce diagramme est appelé dendogramme. Il montre comment on peut choisir un nombre donné de grappes en coupant « l'arbre » par une ligne horizontale à n'importe quel niveau. Par exemple, une ligne horizontale tracée

près du haut de l'arbre produit trois grappes, formées par les ensembles (1), (7) et (2, 3, 4, 5, 6, 8). Dans la plupart des situations pratiques, on doit porter des jugements subjectifs pour déterminer le nombre de grappes à former à partir d'une classification hiérarchique. Everitt, Landau et Leese (2001) décrivent des moyens formels de traiter cette question.

Tableau XVIII.2

Données relatives aux exploitations indiquant la présence ou l'absence d'un ensemble de caractéristiques communes

Caractéristiques	Exploitation (exploitant)							
	1	2	3	4	5	6	7	8
Montagne (+)/bas-fonds (-) ?	-	+	+	+	+	+	-	+
Fortes précipitations ?	-	+	+	+	+	-	-	+
Revenu élevé ?	-	+	+	-	-	+	-	-
Famille nombreuse (>10 membres) ?	-	+	+	+	-	+	-	+
Accès à du bois de feu à moins de 2 km ?	+	-	-	+	+	-	+	+
Centre de santé à moins de 10 km ?	+	-	-	-	-	-	-	-
Femme chef de famille ?	+	-	-	-	-	-	-	-
Eau courante ?	-	-	-	-	-	-	+	-
Présence de latrines sur l'exploitation ?	+	-	-	-	-	+	-	-
Culture de maïs ?	+	-	-	+	+	-	+	+
Culture de pois cajan ?	-	+	+	+	-	+	-	-
Culture de haricots ?	-	-	-	+	+	-	-	+
Culture d'arachides ?	-	-	-	-	-	-	-	+
Culture de sorgho ?	+	-	-	-	-	-	-	-
Élevage ?	+	+	-	-	+	-	+	-

Tableau XVIII.3

Matrice de similarité entre huit exploitations

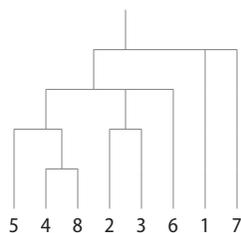
		Exploitation							
		1	2	3	4	5	6	7	8
Exploitation	1	-	1	0	2	3	1	3	2
	2		-	5	4	3	4	1	3
	3			-	4	2	4	0	3
	4				-	5	3	2	6
	5					-	1	3	5
	6						-	0	2
	7							-	2
	8								-

27. Avec un logiciel approprié, l'analyse des grappes peut se faire très simplement, mais seulement après avoir examiné de près les types de données utilisées, le degré de similarité ou de divergence et la méthode employée pour former les grappes. Il faut prendre un soin particulier si le logiciel utilisé ne permet de former de grappes qu'avec un seul type de données. Par exemple, SPSS (2001) exige que toutes les varia-

bles utilisées dans la formation de grappes soient continues, catégoriques ou binaires. Si l'on dispose d'un mélange de types de données, la meilleure solution à employer avec ce type de logiciel peut être de convertir toutes les variables en variables binaires et d'appliquer un degré de similarité convenant à ces variables, tout en admettant, cependant, qu'une telle formule peut entraîner une certaine perte d'informations.

Figure XVIII.2

Dendrogramme formé d'après la matrice de similarité entre exploitations



28. Deux autres questions ne doivent pas être perdues de vue. La première a trait à la nécessité de tenir compte du fait que l'on ignore (à la connaissance de l'auteur) l'impact d'échantillons de conception complexe sur l'analyse des grappes. Si la conception de l'enquête comportait une procédure d'échantillonnage en grappes, avec des différences notables entre les grappes sélectionnées, une analyse des grappes appliquée à l'ensemble des données de l'échantillon sans tenir compte des pondérations de l'échantillonnage pourrait engendrer les grappes de l'enquête elles-mêmes. Il serait donc judicieux d'envisager d'utiliser une analyse des grappes avec chacune des grappes de l'enquête et d'étudier l'homogénéité des résultats obtenus. Là encore, il convient de tenir compte des différentes pondérations de l'échantillonnage présentes à l'intérieur des grappes et d'interpréter les résultats avec prudence si le logiciel ne peut tenir compte des pondérations.

29. La seconde question concerne la possibilité de difficultés de calcul dues aux limites de la mémoire informatique. Ces difficultés peuvent apparaître si l'on effectue l'analyse des grappes en utilisant l'échantillon complet de l'enquête. Pour atténuer les effets de ce problème, si les objectifs de l'analyse des grappes le permettent, l'analyse peut se limiter à des groupes plus restreints de cet échantillon.

F. Analyse en composantes principales (ACP)

30. Supposons que l'on ait plusieurs variables, 12 par exemple, qui mesurent divers aspects de l'un des thèmes principaux de l'enquête. Par exemple, dans une enquête sur la nutrition, l'état nutritionnel des enfants peut se mesurer au moyen de plusieurs mesures anthropométriques, ainsi qu'à l'aide de variables décrivant les caractéristiques socio-économiques de leurs familles. Ces variables ont des chances d'être corrélées, et il se pose alors la question de savoir si l'on peut réduire leur nombre à certaines variables qui illustrent le mieux possible la variation de l'ensemble de données initiales. C'est ce que vise l'analyse en composantes principales (ACP). Cette technique est strictement applicable à un ensemble de mesures soit quantitatives, soit effectuées à l'aide d'une échelle ordinale. Toutefois, comme il s'agit essentiellement d'une technique descriptive, l'inclusion de variables binaires et/ou d'un petit nombre de variables catégoriques nominales ne risque guère d'avoir des conséquences pratiques.

31. Avec l'ACP, on est en présence d'un nouvel ensemble de variables créé sous la forme de combinaisons linéaires² de l'ensemble initial. Cette combinaison linéaire qui explique le degré de variation maximale est appelée la première composante principale. On crée ensuite une seconde composante principale (une autre combinaison linéaire), indépendante de la première, qui explique, autant que possible, la variabilité restante. D'autres composantes sont ensuite créées de façon séquentielle, chaque nouvelle composante étant indépendante des précédentes. Si les quelques premières composantes, disons les trois premières, expliquent en grande partie, par exemple à 90%, la variabilité de l'ensemble initial de 12 variables, le nombre de variables à analyser est ainsi ramené de 12 à 3.

32. Il importe de souligner que les estimateurs de la composante principale peuvent être fortement entachés de distorsions si l'on applique l'ACP à l'ensemble de l'échantillon de l'enquête lorsqu'il n'est pas autopondéré (Skinner, Holmes et Smith, 1986). Comme il a été montré à la section B, l'ACP n'est généralement recommandée pour l'analyse des données d'une enquête que pour les sous-ensembles plus petits de l'échantillon qui ont (au moins approximativement) les mêmes pondérations d'échantillonnage. Si le sous-ensemble de données étudié présente des pondérations sensiblement différentes, il faut faire preuve de prudence dans l'interprétation des résultats.

Exemple 3

33. Pomeroy *et al.* (1997) ont appliqué l'ACP aux données d'une enquête qui portait sur 200 ménages, où il avait été demandé aux répondants de noter 10 indicateurs, sur une échelle de 1 à 15, où les degrés étaient présentés comme les barreaux d'une échelle, pour illustrer leur perception des changements survenus par suite de projets de gestion des ressources communautaires exécutés dans leur région. Ces indicateurs sont énumérés ci-dessous, et les résultats de l'ACP sont présentés au tableau XVIII.4.

Tableau XVIII.4
Résultats d'une analyse en composantes principales

Variable	Composante		
	PC1	PC2	PC3
1. Situation générale du ménage	0,24	0,11	0,90
2. Situation générale des ressources halieutiques	0,39	0,63	0,02
3. Revenu local	0,34	0,51	0,55
4. Accès aux ressources halieutiques	- 0,25	0,72	0,17
5. Contrôle des ressources	0,57	0,40	0,12
6. Aptitude à participer aux affaires communautaires	0,77	0,13	0,29
7. Aptitude à influencer sur les affaires communautaires	0,75	0,22	0,34
8. Conflit au sein de la communauté	0,78	0,03	0,18
9. Respect des règles par la communauté et gestion des ressources	0,82	0,12	0,07
10. Quantité d'eau traditionnellement prélevée	0,38	0,66	0,12
Pourcentage de variance expliqué	33	19	14

La première composante principale est donc donnée par la formule :

$$PC1 = 0,24 (\text{ménage}) + 0,39 (\text{ressource}) \dots + 0,82 (\text{respect}) + 0,38 (\text{prélèvement}).$$

² Si X_1, X_2, \dots, X_p sont l'ensemble initial de p variables, une variable Y issue d'une combinaison linéaire de ces variables prend la forme $Y = a_1X_1 + a_2X_2 + \dots + a_pX_p$, où les a_i ($i=1,2,\dots,p$) sont des nombres, c'est-à-dire les coefficients de la composante principale.

34. Cette première composante est décrite par Pomeroy *et al.* (1997) comme un indicateur du comportement des membres de la communauté; la deuxième composante, comme se rapportant aux ressources halieutiques; et la troisième composante comme un indicateur du bien-être des ménages. Les auteurs utilisent ensuite ces composantes comme les variables dépendantes dans des analyses de régression multiple pour évaluer l'efficacité d'un certain nombre de facteurs explicatifs comme moyens d'expliquer la variabilité de chaque indicateur.

35. Si l'interprétation de ces variables est raisonnable ici, on peut douter de l'intérêt d'utiliser (par exemple) la première composante principale de la façon calculée ci-dessus pour une analyse plus approfondie. Seules les variables 5, 6, 7, 8 et 9 décrivent le comportement des membres de la communauté et obtiennent des notes élevées pour PC1. Plutôt que d'inclure les 10 variables dans le calcul de la première composante principale, il serait préférable de recalculer une nouvelle variable comme simple résumé des variables de comportement de l'ensemble de données initiales, par exemple en prenant une simple moyenne arithmétique des variables 5, 6, 7, 8 et 9 ou une moyenne pondérée de ces variables dans laquelle le contrôle des ressources (variable 5) est affecté d'une pondération légèrement plus faible que les autres. De même, les variables relatives aux ressources (variables 2, 3, 4 et 10) pourraient être combinées en une variable sommaire, tandis que la variable 1 resterait indépendante. Utilisée de cette manière, l'ACP permet de déterminer comment les 10 indicateurs peuvent être résumés simplement pour donner un nouvel ensemble de mesures significatives à analyser plus avant comme l'ont fait, par exemple, Pomeroy *et al.* (1997) par une analyse de régression afin d'étudier les facteurs influant sur chacune des trois composantes principales.

Exemple 4

36. Le cadre définissant un mode de subsistance durable adopté par le Département du développement international [*Department for International Development* (DFID)] du Royaume-Uni de Grande-Bretagne et d'Irlande du Nord nous offre un autre exemple pratique. Ce cadre considère cinq actifs sur lesquels repose le mode de subsistance, à savoir le capital social, le capital humain, le capital naturel, le capital physique et le capital financier. Une enquête effectuée pour étudier les modes de subsistance des ménages nécessiterait que chacun de ces actifs soit mesuré par rapport à un certain nombre de variables subsidiaires. Par exemple, on peut mesurer le capital social du point de vue de son degré de dépendance à l'égard de réseaux d'appui, du pourcentage du revenu des ménages qui provient des envois de fonds de travailleurs émigrés, du niveau de confiance dans le groupe, du degré de participation au processus de décision, etc.; et le capital physique peut se mesurer au fait de posséder une bicyclette ou une radio, d'avoir accès à l'eau courante, à l'électricité, etc.

37. L'objectif ici est de déterminer une variable unique pour chacun des cinq actifs sur lesquels repose la subsistance. Cela peut se faire directement pour les actifs physiques, par exemple par l'établissement d'une simple moyenne pondérée de réponses binaires correspondant au fait qu'un ménage possède ou non certains articles d'une liste donnée, dont les prix unitaires sont utilisés comme pondérations. En revanche, le capital social ne peut être combiné d'une façon simple car il est beaucoup plus difficile d'affecter des pondérations à des variables décrivant des actifs sociaux. Ici, nous pouvons nous voir contraints d'accepter des pondérations dérivées de données par le biais d'une ACP appliquée à un ensemble de variables sociales. Les résultats peuvent être utilisés pour produire une mesure globale acceptable du capital social, une fois encore à condition de tendre vers une moyenne pondérée simple, une fois que

les pondérations relatives de chaque variable de la première ou de la deuxième composante principale sont connues.

G. Méthodes multivariates de construction d'indices

38. La construction d'indices peut avoir plusieurs significations différentes. Par exemple, dans une étude de la santé, on mesure généralement l'état nutritionnel des enfants en établissant des indices à partir de mesures anthropométriques, telles que le poids ou la taille par rapport à l'âge et le rapport entre le poids et la taille, afin de déterminer respectivement les cas d'insuffisance pondérale, d'arrêt de croissance et d'atrophie.

39. Dans un exemple plus complexe, on peut faire la synthèse des réponses à des questions portant notamment sur l'allaitement maternel, l'utilisation de biberons, la diversité diététique, le nombre de jours où l'enfant a reçu certains types d'aliments au cours des sept jours précédents, ainsi que sur la fréquence des repas, afin de créer un indice de l'alimentation de l'enfant (Ruel et Menon, 2002). Il s'agit là d'un autre type d'indice où le chercheur décide des notes à affecter, de manière à faire en sorte que l'échelle ordinale de chaque variable est telle que les valeurs élevées représentent toujours un « bon » ou un « mauvais » résultat. Lorsqu'on est en présence de variables binaires, par exemple pour la possession d'un certain nombre de biens, on peut utiliser le prix du bien pour donner différentes pondérations à chaque question, comme dans l'exemple 4 (section F) ci-dessus.

40. On peut avoir à faire à un autre type d'indice lorsque l'enquête vise à déterminer les attitudes ou les opinions au sujet, par exemple, de l'accès aux services de santé. Dans ce cas, plusieurs questions peuvent se poser, auxquelles la réponse est une note de 1 à 5, allant de « très médiocre » à « très bon ». Là encore, on peut faire la somme des notes données à toutes les questions pertinentes pour établir un indice reflétant l'opinion qu'ont les familles de la qualité de leurs services de santé.

41. Ici, notre analyse va plus loin, englobant des situations où les données déterminent la forme de l'indice par l'utilisation d'une procédure multivariée. Dans ce cas, l'indice continue à être interprété comme une grandeur unique illustrant les informations recueillies sur plusieurs variables sous la forme d'une mesure composite, qui se présente généralement de la façon suivante :

$$\text{Indice} = a_1X_1 + a_2X_2 + a_3X_3 + \dots + a_pX_p$$

où les termes a_i sont les pondérations à déterminer à partir des données et les termes X_i sont un sous-ensemble approprié de variables p mesurées dans le cadre de l'enquête. Nous présentons deux méthodes par lesquelles les pondérations a_i peuvent être déterminées à partir des données (voir ci-après). Le choix de la méthode la plus appropriée dépend généralement des objectifs visés par la construction de l'indice.

42. La première est fondée sur une approche de la modélisation par régression et la seconde sur une application de l'ACP. Ces méthodes sont examinées pour la construction d'indices utilisés pour mesurer des indicateurs indirects de la richesse ou de la situation socio-économique des ménages dans les pays en développement. Il existe de nombreux ouvrages sur ce sujet, dont on peut trouver un aperçu général dans Davis (2002). Voir également le chapitre XVII de la présente publication, qui offre un examen utile de l'utilisation des données provenant d'enquêtes sur les ménages pour comprendre la pauvreté.

1. Modélisation des dépenses de consommation pour la construction d'un indice de substitution du revenu

43. Henstchel *et al.* (2000) et Elbers, Lanjouw et Lanjouw (2001) ont élaboré une approche de la modélisation des dépenses de consommation en tant qu'indicateur de substitution du revenu. Cette approche consiste à utiliser les données d'une enquête approfondie sur le budget des ménages pour identifier les variables indicatives de la pauvreté. Pour ce faire, on utilise les dépenses de consommation comme variable dépendante dans un modèle de régression linéaire multiple et une série de variables indicatives du niveau des ménages (par exemple, les biens détenus par le ménage, la qualité de son logement, l'accès à divers services, etc.) comme variables explicatives éventuelles (prédicteurs). On utilise le meilleur petit sous-ensemble de variables explicatives permettant d'expliquer la variation maximale de la variable (dépendante) de la réponse pour prédire le niveau des dépenses de consommation. Si les variables explicatives proviennent d'un recensement de la population, on peut alors appliquer le modèle d'équation obtenu aux données du recensement pour prédire le niveau des dépenses de consommation de chaque ménage recensé. Ces données peuvent ensuite servir à dresser des cartes de la pauvreté à l'échelle nationale. Si l'enquête sur le budget des ménages a lieu bien avant la date prévue pour le recensement, on peut identifier l'ensemble approprié de variables prédictrices à partir des données de l'enquête sur le budget et inclure cet ensemble dans le questionnaire du recensement. On trouvera ci-dessous un exemple illustrant cette approche.

Exemple 5

44. En 2001-2002, le Bureau national des statistiques de la République-Unie de Tanzanie a entrepris une Enquête nationale sur le budget des ménages (HBS) qui couvrait environ 22 000 ménages. À partir des renseignements recueillis sur les dépenses des ménages au cours d'une période de 28 jours, on a calculé pour chaque ménage le volume total des dépenses de consommation par adulte au cours de cette période. Un modèle de régression utilisant les données préliminaires résultant de l'enquête a permis d'identifier une série de variables définissant le niveau potentiel des ménages (des ensembles distincts ont été utilisés pour les familles urbaines et rurales) et expliquant le haut degré de variabilité des dépenses de consommation. Ces variables ont été incluses dans un questionnaire utilisé dans un recensement des ménages effectué en trois sites témoins retenus par l'équipe du Projet d'étude de la morbidité et de la mortalité des adultes en poste à Dar es Salaam. Il s'agissait de déterminer un indice des dépenses de consommation à partir des données de l'enquête sur le budget des ménages aux trois sites retenus pour ce projet et d'appliquer cet indice aux ménages couverts par ce projet dans chaque site.

45. On trouvera tous les détails des approches de la modélisation ainsi qu'une évaluation de l'efficacité des modèles dans Abeyasekera et Ward (2002). Nous présentons ici un résumé des résultats enregistrés dans une région rurale (voir tableau XVIII.5) afin de montrer les variables prises en compte dans l'équation du modèle et les pondérations (coefficients de régression) utilisées pour calculer un indice des dépenses de consommation.

46. D'après les résultats du tableau XVIII.5, l'indice permettant de prédire les dépenses de consommation des ménages de la région du Kilimandjaro, en République-Unie de Tanzanie se présente comme suit :

$$\begin{aligned} \text{Indice des dépenses de consommation} = & \\ & 9,79388 + (0,11043 \cdot \text{lampe}) + (0,19950 \cdot \text{sofa}) + (0,12870 \cdot \text{bicyclette}) + (0,11858 \cdot \text{semences}) \\ & + (0,16254 \cdot \text{engrais}) + (0,025824 \cdot \text{étendue de terre}) + (0,088769 \cdot \text{viande}) + (0,076132 \cdot \text{revenu4}) \\ & + (0,13451 \cdot \text{revenu3}) + (0,098303 \cdot \text{revenu2}) + (0,27985 \cdot \text{édu4}) + (0,15878 \cdot \text{édu3}) \\ & - (0,0091977 \cdot \text{édu2}) - (0,0022552 \cdot \text{âge}) + (0,010456 \cdot \text{tailleménage2}) - (0,23902 \cdot \text{tailleménage}) \end{aligned}$$

47. Ce modèle explique 65 % de la variabilité des dépenses de consommation, ce qui est un pourcentage sensiblement élevé, étant donné la complexité des conditions de subsistance des familles rurales. On a jugé de la qualité de cet indice au stade de son élaboration : *a*) en le comparant aux montants réels des dépenses de consommation; et *b*) en considérant son aptitude à déterminer la véritable proportion des ménages vivant en dessous du seuil de pauvreté absolue en République-Unie de Tanzanie. La méthode *a*, qui consistait à tracer la courbe de l'indice par rapport aux montants réels, laisse apparaître une excellente concordance. Elle s'est toutefois montrée moins efficace pour le classement de la population en cinq quintiles de richesse : seuls 46 % des ménages ont été classés dans le bon quintile. La classification par rapport au seuil de pauvreté a été meilleure, classant correctement 87 % des ménages au-dessus ou en dessous de ce seuil.

48. D'autres exemples de l'approche de la modélisation sont donnés dans les dernières sections du chapitre XIX.

Tableau XVIII.5

Variables utilisées (et leurs pondérations) pour la construction d'un indice de prévision des dépenses de consommation dans la région du Kilimandjaro, en République-Unie de Tanzanie

Variable prédictrice	Probabilité de signification	Pondération (coefficient du modèle) [estimation STATA]
Taille du ménage	0,000	- 0,239
Carré de la taille du ménage	0,000	0,0104
Âge du chef de famille (années)	0,038	- 0,00226
Instruction du chef de famille ^a	0,000	0, - 0,00920, 0,159, 0,280
Principale source de revenu ^b	0,017	0, 0,0983, 0,1345, 0,0761
Nombre de jours où de la viande a été servie la semaine précédente	0,000	0,0888
Étendue de terre détenue par le ménage	0,000	0,0258
Engrais ^c	0,000	0,1625
Semences ^c	0,004	0,1186
Possession d'une bicyclette	0,000	0,1287
Possession d'un sofa	0,000	0,1995
Possession d'une lampe	0,001	0,1104
Constante dans l'équation du modèle	0,000	9,794

Taille de l'échantillon = 1 026

R² = 0,651

R² ajusté = 0,646

^a Aucune; primaire; secondaire; tertiaire et au-dessus.

^b Vente de récoltes; vente de bétail; commerce/salaire/traitement; autres sources.

^c Achetés au cours des 12 derniers mois.

2. Analyse en composantes principales (ACP) utilisée pour établir un indice de « richesse »

49. La méthodologie examinée à la section G.1 ci-dessus ne peut s'appliquer que si l'on dispose de données fiables sur les dépenses de consommation – la variable dépendante – issues d'une enquête précédente. Le fait qu'il est difficile de recueillir des données fiables sur les dépenses de consommation, combiné aux coûts élevés d'une telle collecte, a poussé certains chercheurs à recommander l'utilisation d'un indice de pauvreté fondé sur les biens, tiré d'une ACP. La première composante principale est utilisée comme indice de la situation socio-économique, conformément à des travaux de recherche antérieurs d'où il ressort que le rapport biens/consommation est un indice très proche de l'indice recherché (Filmer et Pritchett, 1998). Cependant, il faut faire preuve d'une certaine prudence quand on interprète l'indice fondé sur les biens comme mesure de la pauvreté, car l'efficacité d'un tel indice dépend du choix des biens utilisés et de l'ensemble particulier de données auquel est appliquée l'ACP. L'exemple donné de cette approche par Gwatkin *et al.* (2000) illustre la méthode d'ACP employée pour définir les quintiles de richesse en République-Unie de Tanzanie, qui utilise l'ensemble suivant de variables mixtes fondées sur les biens et liées à la santé :

- Le ménage a l'électricité ou possède une radio, une télévision, un réfrigérateur, une bicyclette, une moto, une voiture (dans chaque cas, le code est 1 = oui et 0 = non);
- Nombre de personnes par chambre à coucher (donnée quantitative);
- Principales sources d'alimentation en eau potable (sept catégories);
- Principal type de toilettes utilisé par les membres de la famille (cinq catégories);
- Principal type de matériau utilisé pour le sol du logement familial (six catégories).

50. Les données utilisées sont tirées du questionnaire de l'Enquête démographique et sanitaire. Des pondérations d'échantillonnage appropriées ont été utilisées pour l'analyse.

51. Les auteurs ont tenu à souligner que leurs travaux n'étaient qu'un premier essai appliqué à tout un pays, mais qu'à l'avenir tout effort visant à examiner les différences au sein de la population par catégorie socio-économique produirait des résultats différents. Selon eux, cela pourrait s'expliquer par l'utilisation d'une autre base que les biens pour déterminer la situation socio-économique, par des erreurs d'échantillonnage, etc. Les différences de niveau de richesse selon les endroits seraient une autre raison plus évidente. En fait, on note des différences dans la définition des quintiles de richesse quand on applique leur méthode aux trois sous-catégories de population de la République-Unie de Tanzanie, à savoir les trois régions visées à la section G.2, en utilisant les données de l'Enquête nationale sur le budget des ménages (tableau XVIII.6). Il est donc prudent de ne pas considérer les résultats de l'ACP comme transposables, ne serait-ce qu'à l'intérieur d'un même pays, à des périodes différentes ou si on les applique aux différentes strates de la population

52. Les chercheurs ont également utilisé la première composante principale d'une analyse en composantes principales comme première indication en vue d'une analyse plus approfondie des données. Par exemple, Ruel et Menon (2002) ont construit un indice socio-économique à partir des ensembles de données de l'Enquête démographique et sanitaire pour classer les ménages en terciles aux fins de contrôler la situation socio-économique dans une analyse de régression multiple effectuée pour

déterminer les facteurs influant sur l'état nutritionnel des enfants. Ils ont effectué des analyses distinctes des populations urbaine et rurale, utilisant sept ensembles de données tirés de cinq pays d'Amérique latine. Les variables utilisées étaient les sources d'approvisionnement en eau, l'assainissement, les matériaux de construction des logements (sols, murs, toiture) et la propriété d'une liste de biens. Les valeurs attribuées à ces variables ont été classées par ordre ascendant (de la plus mauvaise à la meilleure) avant d'être soumises à l'analyse en composantes principales. Seules les variables aux coefficients supérieurs à 0,5 ont été retenues pour l'indice final. Cette approche était raisonnable, le principal objectif ayant été la construction d'un indice pour corriger les différences socio-économiques d'une analyse ultérieure.

Tableau XVIII.6
Points de coupure utilisés pour séparer la population en cinq quintiles de richesse

Quintile de richesse	Dar es Salaam ^a (HBS)	Kilimanjaro ^a (HBS)	Morogoro ^a (HBS)	République-Unie de Tanzanie ^a (HBS)	République-Unie de Tanzanie ^b (DHS)
20 ^e pourcentile	- 1,2993	- 0,8452	- 0,9190	- 1,0317	- 0,5854
40 ^e pourcentile	- 0,7709	- 0,6289	- 0,6180	- 0,5704	- 0,5043
60 ^e pourcentile	- 0,1054	- 0,2459	- 0,3645	- 0,3051	- 0,3329
80 ^e pourcentile	1,1603	0,3239	0,4586	0,4609	0,3761

^a Enquête sur le budget des ménages, 2000-2001.

^b Enquête démographique et sanitaire, 1996.

H. Conclusions

53. Dans ce chapitre, nous avons cherché à présenter l'utilisation de méthodes multivariates pour la construction d'indices, en mettant l'accent sur la nécessité, dans un premier temps, de disposer d'outils de recherche multivariates. Toutefois, l'application de ces méthodes exige beaucoup de soin et demande que l'on porte une grande attention à leur signification et à leurs limites. Par exemple, le succès de l'ACP pour la réduction des variables dépend de l'aptitude à condenser une part importante de la variation entre les données à l'aide de quelques indices, et à donner une interprétation significative à chacun d'entre eux. On a également intérêt à bien songer à l'efficacité de l'ACP si la première composante principale ne rend compte que d'une petite fraction seulement de la variation de l'ensemble complet de variables. Il convient également d'accorder l'attention voulue au bon choix des variables incluses dans le calcul de l'indice en regard des objectifs de l'analyse.

54. L'analyse par grappes se heurte à des difficultés liées au choix d'un degré approprié de similarité ou de divergence et aux décisions concernant la méthode à utiliser pour la formation des grappes. Il faut ici tenir compte de divers facteurs, notamment des types de données utilisées, du mode de calcul et de la solidité de la procédure face à de faibles variations des données.

55. Il convient également de souligner une fois encore que les méthodes décrites dans ce chapitre s'appliquent mieux à des sous-ensembles appropriés de la population lorsque celle-ci peut clairement se subdiviser. Tel est particulièrement le cas si les données à analyser proviennent d'une enquête nationale. Il faut choisir judicieusement les sous-ensembles à utiliser. En effet, différents indices peuvent être établis pour différents sous-ensembles. Toutefois, cela en soi est intéressant et semble indiquer qu'une analyse plus approfondie serait préférable pour les sous-ensembles de population étudiés.

56. Ce chapitre a offert une évaluation des techniques multivariées, en tant qu'outil d'explication et, en particulier, pour la construction d'indices. On dispose aujourd'hui de moyens en matière de logiciels de statistiques d'usage général [par exemple, SPSS (2001), STATA (2003)] qui permettent d'effectuer assez facilement de telles analyses. Les chercheurs sont donc invités à les utiliser pour l'analyse des données d'enquêtes afin d'en tirer le maximum possible d'informations et de faciliter la réalisation des objectifs de ces enquêtes.

REMERCIEMENTS

Je tiens à adresser mes sincères remerciements à mon collègue Ian Wilson et à deux collaborateurs anonymes pour leurs précieux commentaires sur les premières versions de ce chapitre. Mes remerciements vont aussi au Bureau national de statistiques de la République-Unie de Tanzanie pour m'avoir permis d'accéder à ses données pour certains exemples utilisés dans cette étude, et je suis également redevable au Department for International Development (DFID) du Gouvernement du Royaume-Uni de Grande-Bretagne et d'Irlande du Nord des idées qu'il a permis d'apporter à la conception de ce chapitre en finançant de nombreux projets d'enquêtes dans les pays en développement. Toutefois, les informations contenues dans ce chapitre n'engagent que la responsabilité de son auteur et n'implique en aucune façon l'expression d'une opinion quelconque de la part du DFID.

RÉFÉRENCES

- Abeyasekera, S., et P. Ward (2002). *Models for Predicting Expenditure per Adult Equivalent for AMMP Sentinel Surveillance Sites*. Dar es Salaam: Adult Morbidity and Mortality, Ministère de la santé de la République-Unie de Tanzanie. Accessible sur www.ncl.ac.uk/ammp/tools_methods/socio.html.
- Chatfield, C., et A. J. Collins (1980). *Introduction to Multivariate Analysis*. Londres, Chapman and Hall.
- Davis, B. (2002). Is it possible to avoid a lemon? Reflections on choosing a poverty mapping method. Accessible sur http://www.povertymap.net/pub/Pov_mapping_methods_18-9-02.pdf.
- Elbers, C., J. Lanjouw et P. Lanjouw (2001). Welfare in villages and towns: micro-level estimation of poverty and inequality. Mimeo. Vrije Universiteit, Université Yale et Banque mondiale.
- Everitt, B. S., et G. Dunn (2001). *Applied Multivariate Data Analysis*. Londres, Arnold.
- _____, S.
- Landau et M. Leese (2001). *Cluster Analysis*. Londres, Arnold.
- Filmer, D., et L. Pritchett (1998). *Estimating Wealth Effects without Expenditure Data—or Tears: An Application to Educational Enrolments in States of India*. Washington, D.C., Document de recherche de la Banque mondiale, n° 1994.
- GenStat (2002). *GenStat for Windows*, 6^e éd. Oxford, Royaume-Uni, VSN International, Ltd.
- Gower, J. C. (1971). A general coefficient of similarity and some of its properties. *Biometrics*, vol. 27, pp. 857-872.
- Gwatkin, D. R., et al. (2000). *Socio-economic Differences in Health, Nutrition and Population in Tanzania*. Washington, D.C., Groupe thématique de la Banque mondiale sur la santé, la population, la nutrition et la pauvreté. Accessible sur <http://www.worldbank.org/poverty/health/data/tanzania/tanzania.pdf> (consulté le 30 juin 2004).

- Hentschel, J., *et al.* (2000). Combining census and survey data to trace spatial dimensions of poverty: a case study of Ecuador. *The World Bank Economic Review*, vol. 14, n° 1, pp. 147-165.
- Jongman, R. H. G., C. J. F. Ter Braak et O.F.R. Van Tongeren (1995). *Data Analysis in Community and Landscape Ecology*. Cambridge, Royaume-Uni, Cambridge University Press.
- Korn, E. L., et B. I. Graubard (1998). Scatterplots with survey data. *The American Statistician*, vol. 52, n° 1.
- Krzanowski, W. J., et F. H. C. Marriott (1994a). *Multivariate Analysis*, part 1. *Distributions, Ordination and Inference*. Londres, Arnold.
- _____ (1994b). *Multivariate Analysis*, part 2. *Classification, covariance structures and repeated measurements*. Londres, Arnold.
- Manly, B. F. J. (1994). *Multivariate Statistical Methods: A Primer*. 2^e éd. Londres, Chapman and Hall.
- Orr, A., et P. Jere (1999). Identifying smallholder target groups for IPM in southern Malawi. *International Journal of Pest Management*, vol. 45, n° 3, pp. 179-187.
- Pomeroy, R. S., *et al.* (1997). Evaluating factors contributing to the success of community-based coastal resource management: the Central Visayas Region Project-1, Philippines. *Ocean and Coastal Management*, vol. 36, n^{os} 1-3, p. 24.
- Ruel, M. T., *et al.* (1999). *Good Care Practices Can Mitigate the Negative Effects of Poverty and Low Maternal Schooling on Children's Nutritional Status: Evidence from Accra*. Food Consumption and Nutrition Division Discussion Paper, No. 62, Washington, D.C., International Food Policy Research Institute.
- _____, et P. Menon (2002). *Creating a Child Feeding Index Using the Demographic and Health Surveys: an Example from Latin America*. Food Consumption and Nutrition Division Discussion Paper, No. 130. Washington, D.C., Institut international de recherches sur les politiques alimentaires.
- Sahn, D. E., et D. Stifel (2000). Assets as a measure of household welfare in developing countries. Working Paper 00-11. Saint Louis, Missouri, Washington University, Center for Social Development.
- SAS (2001). *SAS Release 8.2*. Cary, Caroline du Nord, SAS Institute, Inc., SAS Publishing.
- Sharma, S. (1996). *Applied Multivariate Techniques*. New York, Wiley and Sons, Inc.
- Skinner, C. J., D. J. Holmes et T. M. F. Smith (1986). The effect of sample design on principal component analysis. *Journal of the American Statistical Association*, vol. 81, n° 395, pp. 789-798.
- SPSS (2001). *SPSS for Windows. Release 11.0*. Chicago, Illinois, LEAD Technologies, Inc.
- STATA (2003). *Intercooled Stata 8.0 for Windows*. College Station, Texas, Stata Corporation.
- Tufte, E. R. (1983). *The Visual Display of Quantitative Information*. Cheshire, Connecticut, Graphics Press.
- Wilson, I. M., et R. D. Stern (2001). *Approaches to the Analysis of Survey Data*. Statistical Guideline Series supporting DFID Natural Resources Projects. Reading, Royaume-Uni, Statistical Services Centre, University of Reading. Accessible sur <http://www.reading.ac.uk/ssc> (consulté le 25 juin 2004).

Chapitre XIX

Analyse statistique des données d'enquêtes

JAMES R. CHROMY
 Research Triangle Institute
 Research Triangle Park,
 Caroline du Nord,
 États-Unis d'Amérique

SAVITRI ABAYASEKERA
 Statistical Services
 Centre University de Reading
 Reading, Royaume-Uni de Grande-Bretagne
 et d'Irlande du Nord

RÉSUMÉ

L'analyse d'une enquête doit tenir compte du fait que les données de cette enquête sont tirées d'unités sélectionnées à partir d'échantillons de conception complexe : il faut utiliser des pondérations pour analyser ces données et il faut calculer les variances des estimations de l'enquête d'une manière qui tienne compte de la complexité de la conception de l'échantillon. Le présent chapitre traite de l'établissement des pondérations et de leur utilisation pour le calcul des estimations de l'enquête et présente une analyse générale de l'estimation de la variance pour les données de l'enquête. Il traite d'abord de ce que l'on appelle les estimations « descriptives », telles que les totaux, les moyennes et les proportions, qui sont largement utilisées dans les rapports d'enquête. Puis il examine trois formes d'approches « analytiques » des données d'enquête qui peuvent être utilisées pour examiner les relations entre les variables de l'enquête, à savoir les modèles de régression linéaire multiple, les modèles de régression logistique et les modèles à niveaux multiples. Ces modèles forment un ensemble d'outils précieux pour l'analyse des relations entre une variable de réponse clé et un certain nombre d'autres facteurs. Dans ce chapitre, nous donnons des exemples illustrant l'utilisation de ces techniques de modélisation et des conseils pour l'interprétation des résultats.

Termes clés : conception d'enquête complexe, statistiques analytiques, régression, régression logistique, structures hiérarchiques, modélisation à niveaux multiples.

A. Introduction

1. Les enquêtes sur les ménages utilisent des échantillons de conception complexe pour limiter les coûts de l'enquête. En général, on ne dispose pas de cadres d'échantillonnage complets énumérant tous les individus ou tous les ménages, et même lorsqu'il existe des registres de population, le coût d'exécution d'une enquête par entrevues sur les ménages sur la base d'un simple échantillon aléatoire serait prohibitif. Les enquêtes pour l'étude de la mesure des niveaux de vie (LSMS) examinées au chapitre XXIII offrent un bon exemple de la complexité de conception des enquêtes sur les ménages

2. Le tableau XIX.1 présente un exemple type de structure d'enquête sur les ménages. La plupart des échantillons utilisés pour ces enquêtes sont complexes et im-

pliquent une stratification, un échantillonnage à phases multiples et des taux d'échantillonnage inégaux, comme on l'a vu précédemment. Il faut des pondérations pour compenser l'inégalité des taux d'échantillonnage, et les ajustements effectués pour tenir compte des cas de non-réponse accentuent encore cette inégalité. Il faut tenir compte de la complexité de conception de l'échantillon lorsque l'on estime la précision des estimations de l'enquête.

Tableau XIX.1
Structure type de la conception d'une enquête sur les ménages

Éléments	Définitions possibles	Implications
Strates	Régions Communauté (urbaine ou rurale)	Peuvent réduire les erreurs types des estimations. Le contrôle de la distribution de l'échantillon peut aboutir à un échantillonnage disproportionné.
Unités d'échantillonnage de la première phase	Zones d'énumération du recensement ou zone géographiques similaires Villages des strates rurales	Facilitent la mise en grappes de l'échantillon pour limiter les coûts. Facilite l'établissement de cadres complets d'habitation uniquement dans les zones sélectionnées. Sélection selon une probabilité proportionnelle à la taille.
Unités d'échantillonnage de la deuxième phase	Adresses des unités d'habitation	Peuvent n'avoir aucun ménage, avoir un ménage ou plus ou une personne étrangère. Sélectionnées selon la même probabilité que les unités d'échantillonnage de la première phase.
Unités d'échantillonnage de la troisième phase (lorsque tous les membres de la famille ne sont pas automatiquement inclus dans l'échantillon)	Membres de la famille	Échantillon sélectionné à partir d'une liste des membres des familles obtenue auprès d'un membre de la famille adulte responsable. Peut engendrer des inégalités de pondération pour tenir compte de la taille du ménage.
Unités témoins	Ménages Membres de familles Entreprises agricoles ou commerciales exploitées par les membres de la famille Dossiers spéciaux pour certains sous-groupes, par exemple les adultes actifs. Événements ou incidents touchant des membres de la famille Mesures répétées dans le temps (enquêtes par panel)	Peut demander davantage qu'un fichier analytique pour des analyses à but particulier.

3. La section B du présent chapitre traite de l'établissement de pondérations à utiliser dans l'analyse de l'enquête et de l'utilisation de pondérations pour la production de simples estimations « descriptives », telles que les totaux, les moyennes et les proportions/pourcentages, qui sont amplement utilisées dans les rapports d'enquête. Elle donne aussi un aperçu de l'estimation de la variance pour les estimations fondées sur des échantillons de conception complexe.

4. Les autres sections portent sur trois formes d'utilisation « analytique » des données d'enquêtes qui permettent d'étudier la façon dont une réponse clé ou une variable dépendante – par exemple, les résultats scolaires d'une élève, le niveau de pauvreté d'un ménage – varie en fonction d'un certain nombre de facteurs, souvent appelés variables explicatives ou variables de régression. Les modèles de régression linéaire multiple conviennent lorsque la réponse clé est une variable de mesure quantitative, tandis que les modèles de régression logistique se prêtent mieux aux cas où la réponse clé est une variable binaire, c'est-à-dire que les réponses ne peuvent prendre que deux valeurs possibles (par exemple, oui/non ou présent/absent). Ces méthodes de régression peuvent s'appliquer à un ensemble non emboîté de données d'enquête ou à des unités d'échantillonnage appartenant à un seul niveau de la hiérarchie d'un échantillon à phases multiples. Inversement, l'analyse peut avoir à tenir compte des diffé-

rentes sources de variabilité apparaissant aux différents niveaux hiérarchiques, et dans ce cas, la modélisation à niveaux multiples entre en jeu. Cette approche tient compte de la corrélation entre les unités d'échantillonnage d'un même niveau car ces unités se situent à différents niveaux.

B. Statistiques descriptives : pondérations et estimation de la variance

5. Les enquêtes sur les ménages sont généralement conçues pour produire des estimations des totaux de population, des moyennes de population ou de ratios simples de ces totaux ou de ces moyennes. Comme exemples de totaux, on peut avoir la population totale, le nombre d'hommes appartenant à la population active, le nombre de femmes appartenant à la population active, le nombre d'enfants âgés de cinq ans ou moins. Comme exemples de moyennes, on peut citer le revenu moyen des personnes appartenant à la population active, le revenu moyen des femmes appartenant à la population active, ou le revenu moyen des hommes appartenant à la population active. On peut avoir besoin d'estimations de ratios pour estimer la proportion de ménages dont le revenu total se situe en dessous du seuil de pauvreté ou le revenu moyen des ménages dont le principal soutien est une femme.

6. Les enquêtes sur les ménages produisent des estimations nationales, mais peuvent aussi être conçues pour produire des estimations correspondant à des régions géopolitiques ou d'autres estimations transversales. Par ailleurs, ces enquêtes peuvent être répétées pour donner des estimations périodiques (par exemple, annuelles ou quinquennales) qui peuvent être perçues comme des études temporelles. Tant que les statistiques produites ne sont que des estimations de totaux, de moyennes ou de taux, même si elles portent sur une section en coupe de la population ou sur une période donnée, nous qualifions l'analyse nécessaire pour produire ces estimations de « descriptive ». Les statistiques descriptives comprennent les estimations elles-mêmes, ainsi qu'une certaine mesure de la précision de ces estimations. Les rapports descriptifs peuvent inclure les erreurs types des estimations ou les estimations des intervalles fondées sur ces erreurs types. L'estimation des erreurs types nécessite une analyse qui tienne compte de la conception de l'échantillon de l'enquête sur les ménages. Les estimations des intervalles nécessitent non seulement les estimations appropriées des erreurs types mais aussi la connaissance du degré de liberté dans le calcul des estimations de ces erreurs. Ces types de statistiques descriptives relativement simples constituent la majorité des statistiques officielles publiées pour décrire les résultats des enquêtes sur les ménages.

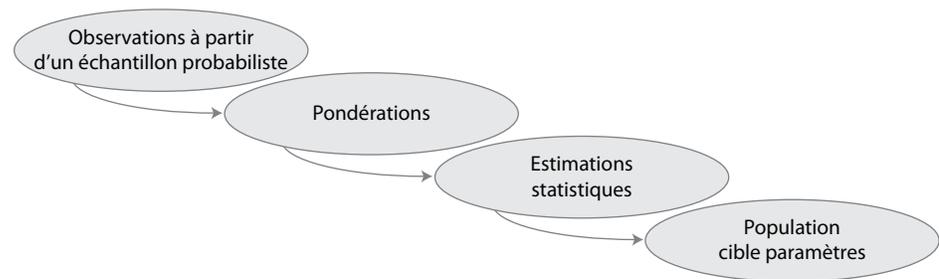
7. Les pondérations¹ de l'enquête et l'estimation statistique fondée sur ces pondérations fournissent le lien entre les observations tirées d'un échantillon probabiliste de ménages et les mesures sommaires ou les paramètres de population applicables aux ménages. La figure XIX.1 illustre ce lien. La population correspondant à tous les ménages est parfois appelée population cible ou univers. Sans l'application de l'échantillonnage probabiliste et de la pondération, aucune théorie statistique ne fournit de lien entre les observations tirées de l'échantillon et les paramètres de la population cible.

8. Toute analyse qui ne tient pas compte de la conception de l'échantillon et des pondérations doit reposer sur des hypothèses. Si l'échantillon est conçu pour engendrer un échantillon probabiliste équivalent, on peut sans danger ne pas faire cas des pondérations pour l'estimation des moyennes, des taux ou des relations entre les variables. Kish (1965, pp. 20-21) appelle ces conceptions *epsem* (*equal probability of selection method sampling*, ou échantillonnage à égale probabilité de méthode de sélection)

¹ Les pondérations fondées sur la conception sont généralement définies comme l'inverse de la probabilité de sélection des unités d'observation sélectionnées. Les pondérations de l'enquête qui figurent dans les dossiers de l'analyse pour les enquêtes sur les ménages sont généralement fondées sur la conception et ont été ajustées pour tenir compte des cas de non-réponse. Souvent, des ajustements supplémentaires sont apportés pour permettre la post-stratification ou pour faire en sorte que l'échantillonnage concorde avec des totaux marginaux connus ou beaucoup plus précis. En outre, on peut recourir à une certaine forme d'allègement des pondérations pour limiter l'effet de pondérations inégales lorsque de fortes pondérations sont dues à un échantillonnage imprévu ou à des collectes de données sur le terrain. On utilise l'expression « pondérations de l'enquête » pour différencier ces pondérations des pondérations strictement fondées sur la conception.

et note que même des échantillons complexes à phases multiples peuvent être conçus pour être *epsem* pour les unités d'échantillonnage au stade final ou presque final de la conception. Comme on le verra plus avant, les ajustements pour tenir compte des cas de non-réponse pourraient créer des pondérations inégales même si la conception initiale est *epsem*. Si l'on applique l'étalonnage post-stratification ou multidimensionnel aux données en ajustant les pondérations, on obtient presque toujours des ajustements de pondérations inégaux et, par conséquent, des pondérations inégales.

Figure XIX.1

Application de pondérations et estimation statistique

9. Toutefois, certains analystes sont prêts à formuler des hypothèses qui devraient permettre d'analyser les données des enquêtes sur les ménages sans pondérations ou avec des pondérations égales. Ces hypothèses sont particulièrement valables lorsque l'on applique des modèles aux données pour étudier les relations entre une variable dépendante et un certain nombre de variables explicatives.

10. Pour le cas théorique d'enquêtes où tous les membres de l'échantillon donnent une réponse complète, l'utilisation de pondérations fondées sur la conception calculées comme l'inverse de probabilité de sélection de chaque unité d'observation produit des estimations non faussées des totaux de population et autres statistiques linéaires (Horvitz et Thompson, 1952). Dans la pratique, les enquêtes sur les ménages se heurtent toujours à des cas de non-réponse qui peuvent entraîner une distorsion des estimations si l'on retire ces observations de l'analyse sans prendre d'autres dispositions (voir chapitre VIII). Des techniques ont été mises au point pour tenter de réduire la distorsion due aux cas de non-réponse. L'approche la plus simple consiste à fractionner l'échantillon de manière qu'à l'intérieur de chaque catégorie les différences entre les paramètres de population applicables aux répondants et aux non-répondants puissent être considérées comme très faibles ou insignifiantes (Rubin, 1987). Les ajustements des ratios aux pondérations sont alors effectués dans les catégories de pondérations de manière que chaque catégorie soit représentée dans les estimations ajustées dans la même proportion que celle dans laquelle elle aurait été représentée dans l'échantillon sélectionné.

11. Le processus d'échantillonnage probabiliste ne garantit pas nécessairement que la distribution de l'échantillon sélectionné selon des caractéristiques connues sera identique à celle de la population totale. La stratification avant la sélection de l'échantillon permet d'assurer qu'il en sera ainsi pour certaines caractéristiques, mais cela peut être impossible pour d'autres si la variable de classification ne figure pas sur le cadre utilisé pour sélectionner l'échantillon. Au lieu d'avoir à effectuer de complexes ajustements de ratio pour chaque estimation tirée des données de l'enquête sur les ménages, on introduit souvent un ajustement unique post-stratification des pondé-

rations, que l'on applique ensuite automatiquement à toutes les estimations produites en utilisant les pondérations ajustées. L'approche la plus simple de l'ajustement post-stratification utilise un fractionnement de l'échantillon similaire à celui utilisé pour l'ajustement des pondérations pour les cas de non-réponse.

12. On utilise parfois les méthodes de calibration qui contrôlent la distribution de l'échantillon pondéré simultanément en plusieurs dimensions pour ajuster les pondérations pour les cas de non-réponse, pour l'ajustement post-stratification ou pour les deux (Deville et Särndal, 1992; Folsom et Singh, 2000).

13. Des pondérations extrêmement élevées peuvent accentuer la variance des estimations tirées des enquêtes sur les ménages par un effet de conception (voir chap. VI et VII). Parfois, on réduit arbitrairement ces pondérations, surtout si elles ne sont pas le résultat de la conception prévue de l'échantillon.

14. Les pondérations finales affectées à un fichier analytique produit à partir d'une enquête sur les ménages peuvent comprendre les éléments suivants :

- La pondération fondée sur la conception, calculée comme l'inverse de la probabilité globale de sélection;
- Un élément d'ajustement pour les cas de non-réponse;
- Un élément d'ajustement post-stratification;
- Un élément d'allègement des pondérations.

15. Ces éléments ont besoin d'être décrits pour que l'analyste puisse les examiner. Les éléments d'ajustement appliqués aux pondérations initiales fondées sur la conception impliquent une part de jugement subjectif et parfois arbitraire dans la définition des catégories de pondération, dans le choix des totaux de contrôle pour l'ajustement post-stratification et dans le degré d'allègement appliqué pour contrôler l'effet lié à la conception. Lorsque les estimations de l'enquête laissent apparaître des résultats inattendus ou des anomalies, il peut être bon de procéder à un examen minutieux du processus de pondération ainsi que des autres aspects de la conception et de l'exécution de l'enquête.

16. En général, l'utilisation analytique des données des enquêtes sur les ménages présente des difficultés particulières dues à la complexité de la conception des enquêtes, qui prévoit notamment l'utilisation de pondérations et d'une certaine structure. Les effets dus à la complexité de conception de l'enquête sont examinés dans plusieurs chapitres du présent ouvrage. Le chapitre XX, en particulier, traite de l'impact de cette complexité sur les résultats de l'analyse. Pour un examen plus approfondi de l'analyse d'une enquête complexe ou pour plus de détails sur certains sujets, il est recommandé au lecteur de se reporter à Skinner, Holt et Smith (1989); Korn et Graubard (1999); et Chambers et Skinner (2003). Le chapitre XX de cette publication contient également un examen plus technique de l'analyse d'enquêtes complexes, et le chapitre XXI examine le logiciel et offre des exemples d'approches de l'analyse de données d'enquêtes à partir d'exemples réels.

17. **Statistiques non linéaires.** Même des statistiques simples, telles que les moyennes, deviennent non linéaires dans les enquêtes complexes. Pour estimer une moyenne de population à partir d'une enquête complexe, il faut estimer la population totale à laquelle s'applique la variable étudiée, par exemple le revenu familial, et estimer la taille de la population, disons le nombre total de familles. La moyenne est alors estimée comme le ratio entre ces deux estimations. Le revenu moyen des familles serait estimé comme :

$$\text{Estimation du revenu moyen des familles} = \frac{\text{Estimation du revenu familial total}}{\text{Estimation du nombre total des familles}}$$

Cette moyenne estimative est en fait une fonction non linéaire (ratio) de deux statistiques linéaires. Dans les enquêtes complexes, la taille de l'échantillon (nombre d'observations d'un type particulier) est en soi une variable aléatoire. Ces types d'estimations non linéaires ne sont pas entachés de distorsions pour de petits échantillons, et sont homogènes en ce sens que si l'on étend la taille de l'échantillon à l'ensemble de la population étudiée, l'estimation non linéaire est exactement égale au chiffre de la population en question (Cochran, 1977, pp. 21, 153 et 190). Si nous considérons la population finie comme découlant d'une population infinie hypothétique, nous pouvons envisager de laisser la taille de l'échantillon s'accroître de façon illimitée. Dans ce cas, nous pouvons prétendre à l'homogénéité du modèle lorsque l'estimation non linéaire converge en probabilité vers le paramètre de surpopulation à mesure que s'accroît l'échantillon (voir, par exemple, Skinner, Holt et Smith, 1989, pp. 17-18).

18. On ne peut indiquer qu'approximativement les erreurs types des statistiques non linéaires en utilisant les approximations de premier ordre de la série de Taylor. On peut obtenir des estimations des erreurs types de statistiques non linéaires en utilisant les approximations de premier ordre de la série de Taylor ou des méthodes de réplification telles que la réplification répétée équilibrée ou « jackknife replication ».

19. On peut appliquer les mêmes arguments à l'analyse utilisant des modèles « linéaire » lorsque les fonctions linéaires de la variable dépendante et de la variable indépendante sont estimées d'abord pour l'ensemble de la population.

20. En résumé, l'utilisation de pondérations engendre des estimations linéaires non faussées et conformes aux estimations non linéaires. Dans la pratique, l'utilisation d'estimations homogènes est considérée comme un moyen satisfaisant de contrôler la distorsion des estimations. D'autres types de distorsions et d'erreurs autres que les erreurs d'échantillonnage, telles que celles qui découlent des cas de non-réponse, d'erreurs de la part de l'enquêteur ou du répondant, sont généralement beaucoup plus significatives, surtout sur les échantillons de grande taille.

21. **Structure de l'échantillon dans les enquêtes sur les ménages.** En général, la population et l'échantillon peuvent tous deux avoir une certaine structure. Dans les échantillons des enquêtes sur les ménages, une structure emboîtée est généralement imposée au cadre d'échantillonnage, comme on l'a vu dans la section précédente et comme le montre le tableau XIX.1. Si la structure n'influe pas sur l'établissement d'estimations statistiques de premier ordre telles que totaux, moyennes, ratios ou coefficients de modèles, elle a un effet sur les statistiques de second ordre (estimations de variance), ce qui permet aux analystes d'estimer les erreurs types des statistiques de premier ordre et de concevoir des tests de signification statistique de certaines hypothèses.

22. L'expression complète de la variance d'estimations fondées sur des échantillons stratifiés à phases multiples comprend des composantes pour chaque phase de l'échantillon. Par exemple, si l'on n'emploie la stratification qu'à la première phase, une estimation \hat{T} d'une population totale T fondée sur un échantillon à trois phases avec des segments correspondant à des endroits, aux ménages et aux membres des ménages pourrait avoir une variance exprimée sous la forme suivante :

$$\text{Var}(\hat{T}) = \sum_h \left(fpc_{h1} \frac{S_{h1}^2}{n_{h1}} + fpc_{h2} \frac{S_{h2}^2}{n_{h2}} + fpc_{h3} \frac{S_{h3}^2}{n_{h3}} \right)$$

où les termes de la strate h sont définis comme suit : les termes fp_{chi} sont les facteurs de correction de la population finie au stade de la sélection du segment correspondant à l'endroit ($i = 1$), au stade de la sélection des unités d'habitation ($i = 2$), et au stade de la sélection des personnes ($i = 3$). Les termes S_{hi}^2 sont les composantes de la variance fondées sur les données pondérées aux trois phases de l'échantillonnage. Les termes n_{hi} sont les tailles des segments ($i = 1$), des ménages ($i = 2$) et le nombre de personnes ($i = 3$) de la strate h . Dans la pratique, il n'est pas rare que certaines de ces composantes de la variance soient difficiles, voir impossibles à estimer; cela peut être dû à des sous-échantillons de taille 1 ou à d'autres raisons. Cochran (1977, p. 279) note que, si l'on peut ne pas tenir compte du coefficient de correction de la population finie à la première phase (que l'on suppose être 1), on peut fonder les estimations de la variance sur un équivalent de cette formule beaucoup plus simple qui ne fait intervenir que la première phase de l'échantillonnage. L'hypothèse d'un coefficient de population finie de première phase égal à l'unité est souvent décrite comme estimation de la variance de l'échantillon « avec remplacement » proche de la variance d'un échantillon « sans remplacement ».

23. Pour faire en sorte que cette approche puisse s'appliquer aux estimations linéaires des totaux de population pour l'échantillon à trois phases évoqué ci-dessus lorsque les unités d'observation sont les personnes, on peut définir une nouvelle variable :

$$Z_{hi} = n_{h1} \sum_j \sum_k w_{hijk} Y_{hijk}$$

où w_{hijk} et Y_{hijk} sont respectivement la pondération et la variable observées pour la personne k du ménage j de l'endroit i dans la strate h . Dès lors, on peut obtenir une estimation raisonnable de la variance sous la forme suivante :

$$\text{var}(\hat{T}) = \sum_h \frac{\sum_i (Z_{hi} - \bar{Z}_h)^2}{n_{h1}(n_{h1} - 1)}$$

Cela est possible avec cette formulation car l'estimation de la population totale peut être définie comme :

$$\hat{T} = \sum_h \bar{Z}_h$$

Si l'on fait un choix approprié de Z_{hi} , on peut estimer les variances des statistiques linéaires et non linéaires en utilisant les approximations de premier ordre de la série de Taylor². On peut également estimer les paramètres de la régression ou de la régression logistique. Les contributions de la variance des phases suivantes n'ont pas besoin d'être estimées pour que cette approche fonctionne.

24. Si la correction de la population finie de la première phase est sensiblement inférieure à 1, cette formulation entraîne une surestimation de la variance et de l'erreur type des estimations de l'enquête. Une légère surestimation pourrait déboucher sur des intervalles de confiance relativement larges ou sur moins de déclarations statistiquement significatives lors de l'exécution de tests sur les hypothèses. À ce titre, la supposition d'une correction de la population finie de la première phase égale à 1 est jugée statistiquement prudente, car elle permet de se protéger de fausses déclarations statistiquement significatives. Il convient de noter que l'application du logiciel fondé sur la série de Taylor et du logiciel fondé sur la réplication est simplifiée par l'hypo-

² Woodruff (1971) montre comment on peut concevoir des variables linéarisées pour faciliter le calcul d'approximations complexes de la variance dans la série de Taylor.

thèse d'un facteur de correction de la population finie de 1 lors de la première phase de l'échantillonnage (voir chap. XXI).

C. Statistiques analytiques

25. Dans la présente section, nous passons de l'examen de simples estimations descriptives à ce que l'on appelle des « statistiques analytiques », c'est-à-dire des statistiques qui examinent les relations entre les variables. En fait, dès que les utilisateurs des données veulent comparer des estimations entre divers domaines, les statistiques deviennent « analytiques ». Les statistiques analytiques simples peuvent consister à comparer les différences entre domaines, comme lorsque l'on compare, par exemple, la proportion de ménages dont le revenu total se situe en dessous du seuil de pauvreté dans deux subdivisions géopolitiques ou la production agricole des deux dernières années. Parfois, dans une simple comparaison, les estimations sont indépendantes l'une de l'autre, de sorte que l'erreur type (*se*) de la différence peut être déterminée strictement à partir de l'erreur type des deux estimations. Dans ce cas, on peut déterminer l'erreur type de la différence estimée entre deux domaines à l'aide de l'équation suivante :

$$se(\bar{y}_1 - \bar{y}_2) = \sqrt{\{se(\bar{y}_1)\}^2 + \{se(\bar{y}_2)\}^2}$$

Cette formule de détermination de l'erreur type d'une différence suppose que les deux estimations sont indépendantes et que, de ce fait, elles sont sans corrélation. Cette forme d'erreur type sur des différences est commode pour les utilisateurs des données, car ils peuvent déterminer l'erreur type d'une différence à partir d'erreurs types publiées des différentes estimations. Toutefois, dans le cas d'échantillons complexes, les estimations des domaines sont souvent corrélées. La variance de la différence entre les estimations de deux domaines comporte alors un terme de covariance :

$$se(\bar{y}_1 - \bar{y}_2) = \sqrt{\{se(\bar{y}_1)\}^2 + \{se(\bar{y}_2)\}^2 - 2\text{cov}(\bar{y}_1, \bar{y}_2)}$$

26. Ce terme de covariance est généralement positif, et se traduit donc par une erreur type plus faible sur l'estimation de la différence que dans le cas d'estimations indépendantes examiné ci-dessus. On peut concevoir les enquêtes sur les ménages de manière à tirer profit de ce terme de covariance dans les erreurs types des estimations de différences; les enquêtes longitudinales par panel produisent une covariance fortement positive entre les estimations annuelles en utilisant un échantillon commun continue de personnes ou de ménages. Étant donné que l'on ne peut déterminer l'erreur type de la différence à partir des erreurs types publiées de différentes estimations, il devient nécessaire de prévoir quelles sont les comparaisons qui présentent le plus d'intérêt et de publier également leurs erreurs types.

27. Pour les statistiques strictement descriptives de populations finies, on peut réduire correctement l'erreur type d'estimations descriptives en appliquant un coefficient de correction de la population finie (*fpc*). Dans le cas le plus simple d'un simple échantillonnage aléatoire, ce coefficient est donné par l'équation :

$$fpc = 1 - \frac{n}{N}$$

où *n* est la taille de l'échantillon et *N* celle de la population. Si l'objet de l'analyse est purement analytique, même dans le cas le plus simple statistiquement significatif de

la différence observée entre les moyennes de deux domaines, on ne peut utiliser le coefficient de correction de la population finie (Cochran, 1977, pp. 34-35), parce que la forme du test de signification statistique exige que l'on décide si les populations des deux domaines ont pu provenir d'une même population hypothétique infinie (une seule super-population)³. L'utilisation de coefficients de correction de la population finie dans une structure complexe est examinée plus avant.

D. Commentaires généraux sur la modélisation par régression

28. Les méthodes examinées dans les autres sections de ce chapitre portent sur une technique de modélisation qui établit un modèle de la variation d'une variable de réponse clé ou d'une variable dépendante, et détermine quel sous-ensemble d'un ensemble de variables explicatives potentielles contribue le plus sensiblement à cette variation. On peut choisir ce meilleur sous-ensemble en appliquant les procédures appropriées pour la sélection des variables ou en utilisant une procédure séquentielle pour examiner un certain nombre de modèles différents, en tenant dûment compte du bon choix, d'un point de vue pratique, des variables qui entrent dans ce modèle ou en sont retirées à chaque étape de la procédure analytique.

29. Nous tenons à préciser que les techniques décrites dans ce chapitre sont à considérer comme devant venir s'ajouter, plutôt que se substituer, aux méthodes d'analyse plus simples. Un premier examen des données sous des formes descriptives sommaires simples (moyennes, écarts types, etc.) et de tabulations pertinentes est très utile et devrait constituer la première phase de l'analyse des données. Parfois, cela peut suffire. Souvent, toutefois, les objectifs de l'enquête exigent une analyse des données, auquel cas les techniques de modélisation risquent fort de prendre de l'importance.

30. Les méthodes de modélisation examinées ici sont particulièrement pertinentes lorsque l'approche est globale, par exemple lorsque l'objectif analytique est de comprendre la logique des systèmes agricoles existants et la façon dont les ménages gèrent leurs ressources limitées de manière à répondre à la fois à leurs besoins de production et de consommation. L'important est toujours de veiller à une application pratique d'une technique de modélisation appropriée, sans perdre de vue les difficultés qui peuvent apparaître sur le terrain dans les pays en développement. Nous insistons sur les limitations de l'analyse pour faire en sorte que les approches examinées ne soient appliquées qu'après avoir dûment mesuré le bien-fondé de la méthode appliquée à la recherche envisagée.

31. On utilise les modèles de régression pour mieux comprendre la relation entre une variable dépendante et un ensemble de variables indépendantes ou explicatives. Il convient toutefois de souligner qu'il est généralement impossible d'établir une relation de cause à effet à tout lien observé entre une variable dépendante et une variable explicative, sauf dans le cas d'expériences bien conçues, contrôlées et rendues aléatoires⁴. En gardant cela présent à l'esprit, on peut apprendre beaucoup de l'application des modèles de régression aux données provenant d'enquêtes sur les ménages.

32. Contrairement aux données découlant d'expériences contrôlées qui emploient la simulation de distribution aléatoire et le contrôle des variables auxiliaires, les données des enquêtes sur les ménages sont généralement issues d'observations et n'ont que peu ou pas d'influence sur les autres facteurs pouvant modifier les relations entre variables. Les méthodes de régression peuvent parfois éliminer les effets de ces variables parasites incontrôlées, ce qui permet d'obtenir des estimations moins tendancieuses de la vraie relation.

³ Cochran (1977, p. 39) affirme que l'utilisation d'un coefficient de correction d'une population finie est inacceptable pour tester les différences entre moyennes de domaines. L'interprétation de cette affirmation devient plus ambiguë lorsqu'on l'applique à des échantillons complexes comportant stratification et mise en grappes; Chromy (1998) examine le problème dans le cas d'un échantillonnage d'élèves d'écoles où les écoles sont fortement stratifiées et sélectionnées. Graubard et Korn (2002) présentent un examen récent de cette question.

⁴ Les expériences rendues aléatoires peuvent être intégrées aux enquêtes. Souvent, il s'agit d'expériences méthodologiques d'un échantillon préalable ou d'échantillons supplémentaires pour une enquête en cours. On peut également effectuer des expériences sociales en recrutant des sujets pour une expérience sociale à partir de l'échantillon d'une enquête sur les ménages.

33. La modélisation par régression est souvent de caractère exploratoire. On peut élaborer un certain nombre de modèles différents pour expliquer le comportement d'une variable dépendante. Les variables explicatives utilisées dans le modèle se limitent à celles qui figurent dans le fichier de données de l'enquête; de ce fait, les variables choisies pour expliquer la variation d'une variable dépendante peuvent n'être que de solides corrélats du véritable facteur causatif. Il peut y avoir plusieurs corrélats antagoniques du facteur causatif, dont aucun ne semble logiquement avoir de rapport avec la variable dépendante. Les analystes doivent se laisser guider par la théorie substantive (par exemple, sociale ou économique) dans le choix des variables explicatives et l'établissement du type de relation (par exemple, linéaire ou non linéaire).

34. Lorsque la théorie substantive ne laisse pas entrevoir de fortes relations théoriques ou lorsque plusieurs variables explicatives antagoniques sont suggérées par la théorie substantive, on peut appliquer des approches de la sélection des variables à partir d'ensembles standard (étrangers à l'enquête) pour identifier les éventuelles variables explicatives. Nombre de logiciels conçus à d'autres fins que pour des enquêtes offrent des approches directes et inverses de la sélection des variables qui permettent d'identifier les variables explicatives ayant des relations linéaires avec la variable dépendante. Si ces logiciels le permettent, l'utilisation des pondérations de l'enquête pour cette analyse exploratoire est fortement recommandée. Ces pondérations peuvent être normalisées de manière à permettre de meilleures estimations de l'erreur et des tests plus précis de la signification statistique (voir les exemples de cette approche au chapitre XXI). Après avoir utilisé de tels logiciels ou programmes pour sélectionner les variables, il est bon d'évaluer le modèle à l'aide d'un logiciel qui utilise les pondérations de l'enquête et reconnaît la conception des enquêtes sur les ménages.

35. Les variables peuvent être des variables catégoriques, des variables de comptage ou des variables de mesure continues. On utilise les modèles de régression linéaire lorsque les variables dépendantes sont des mesures de comptage ou continues; les transformations logarithmiques sont recommandées pour les données de comptage. Lorsque la variable de comptage dépendante comprend des valeurs égales à zéro, la transformation logarithmique échoue, mais des procédures telles que la procédure PROC LOGLINK (SUDAAN 2001) peuvent être utilisées pour satisfaire la valeur attendue du logarithme d'une variable de comptage. On utilise la régression logistique lorsque la variable dépendante est une variable catégorique définie à deux niveaux; on peut également appliquer les modèles de régression multinominale aux variables dépendantes catégoriques à plus de deux niveaux. Pour les besoins de l'analyse, on classe les variables explicatives comme catégoriques ou continues, car les variables de comptage et les variables de mesure continues sont traitées essentiellement de la même manière dans le contexte de la modélisation. Les données d'enquêtes peuvent aussi être analysées au moyen de modèles de survie et d'autres techniques multivariates non traitées dans ce chapitre.

36. L'utilisation de variables explicatives catégoriques, qui définissent les domaines de l'étude, est analogue à la construction de simples comparaisons de domaines sans l'utilisation de modèles. L'utilisation de modèles permet à l'analyste d'effectuer un ajustement pour tenir compte d'autres variables explicatives éventuelles. C'est ce que l'on appelle souvent l'ajustement pour tenir compte de covariates. En l'absence d'un tel ajustement, les coefficients du modèle de régression reproduisent de simples comparaisons de domaines et estiment les différences observées dans la population selon les domaines. Lorsque l'on introduit d'autres variables dans le modèle comme covariates, les coefficients du modèle de régression estiment les différences qui seraient théoriquement observées selon les domaines si l'on maintenait les covariates aux mêmes niveaux dans tous les domaines.

37. On peut également obtenir les coefficients du modèle de régression pour les variables explicatives continues avec ou sans ajustement pour tenir compte d'autres covariates. Les décisions d'ajuster ou non doivent être fonction du but de l'analyse. Les estimations non ajustées décrivent une relation empirique entre les variables dépendantes et explicatives telles qu'elles figurent dans la population. Les estimations ajustées décrivent la même relation si d'autres variables sont maintenues théoriquement constantes. Si les autres variables incluses dans le modèle sont aussi de bons prédicteurs de la variable dépendante, elles peuvent améliorer la précision des valeurs prédites pour des niveaux donnés des prédicteurs clés étudiés. Le choix des méthodes d'analyse doit être fonction du but de l'analyse.

38. Les exemples ci-dessous ne traitent que de modèles simples de variables explicatives continues. Lorsque les variables explicatives sont continues, l'analyste doit étudier la relation entre la variable dépendante et d'éventuelles variables explicatives. De simples graphiques peuvent montrer qu'une relation linéaire ne convient pas pour des variables véritablement liées les unes aux autres. Selon les tracés observés, des termes supplémentaires (quadratiques ou cubiques) peuvent être ajoutés pour mieux saisir la relation. La variable dépendante peut alors avoir des relations linéaires avec une variable explicative, avec son carré, avec son cube ou avec sa valeur à des exposants plus élevés. Les tracés résiduels, une fois incluses certaines des variables explicatives potentielles, peuvent servir à déterminer si d'autres variables ou les exposants plus élevés (carrés ou cubes) des variables incluses peuvent influencer sur le modèle. Pour les variables explicatives ayant une large gamme de valeurs et différents effets sur la variable dépendante d'un bout à l'autre de la gamme, des modèles « spline » qui permettent à la relation de varier à travers certaines subdivisions de la gamme sont souvent utiles. Lorsqu'un échantillon d'une enquête englobe les jeunes, les personnes d'âge moyen et les personnes âgées, on peut souvent afficher les effets de l'âge en utilisant des modèles de régression « spline ».

39. Parmi les autres procédures diagnostiques figurent l'examen de la bonne adaptation des modèles proposés et l'examen de la signification statistique des paramètres de régression pour les variables ajoutées. Les procédures types (autres que celles utilisées pour les enquêtes) peuvent être utilisées avec des données d'enquête pondérées et la régression linéaire. On peut utiliser des tableaux d'imprévus pour évaluer le bien-fondé des modèles de régression logistique. Korn et Graubard (1999, chap. 3) offrent une bonne analyse de l'adaptation des procédures diagnostiques à l'analyse générale des données d'enquêtes.

40. Le développement de modèles de régression fondés sur les données observées fait clairement appel à la notion d'analyse exploratoire de données (Tukey, 1977). Ce type d'analyse peut fournir des renseignements utiles sur les données et la relation entre les variables observées, mais la signification statistique des résultats d'une telle analyse « non prévue » devra être confirmée ou validée par l'étude d'autres données de l'enquête.

E. Modèles de régression linéaire

41. Pour analyser les modèles de régression linéaire (qui font l'objet de la présente section) et les modèles de régression logistique (section F), il est utile de supposer que l'échantillonnage est fait « avec remplacement » à la première phase. Nous supposons également que le fichier analytique d'observation de données comprend des variables d'indice pour les strates, désignées par la lettre h , et pour les unités primaires d'échantillonnage (UPE), désignées par la lettre i . Comme il est indiqué à la section B ci-dessus, il n'est pas nécessaire d'identifier d'autres variables de structure

lorsque nous sommes prêts à utiliser l'hypothèse de conception avec remplacement à la première phase de sélection de l'échantillon. Les implications de l'utilisation d'un échantillon de ménages complexe ne sont prises en compte dans les estimations des coefficients du modèle et leurs erreurs types que si nous utilisons un programme statistique qui tient convenablement compte de la conception de l'enquête sur les ménages, y compris des pondérations analytiques et de la structure de conception (strates et UPE).

42. On peut définir un modèle de régression linéaire utilisant une variable explicative continue et une variable explicative catégorique par l'équation suivante :

Modèle 1

$$y_{hij} = \alpha x_0 + \beta_1 x_{1hij} + \sum_{d=1}^D \gamma_d x_{2dhij} + \varepsilon_{hij}$$

43. Dans le modèle 1, les observations sont représentées par la variable dépendante observée, y_{hij} ; une variable ordonnée à l'origine, x_0 , toujours fixée à 1; une variable explicative continue observée, x_{1hij} ; et un ensemble de variables indicatrices, x_{2dhij} , définissant les niveaux D d'une variable catégorique. Les paramètres du modèle de régression α , β_1 , et γ_d ($d = 1, 2, \dots, D$) sont appelés coefficients de régression et sont estimés par l'analyse. Le terme final du modèle est le terme d'erreur et mesure l'écart par rapport au modèle lié à la j^e observation liée à l' i^e UPE de la h^e strate. C'est un modèle des principaux effets, car il ne comporte pas d'effets d'interaction.

44. Selon le logiciel utilisé, l'ensemble de variables indicatrices peut être défini comme variable unique dans une présentation du modèle; il peut être nécessaire de définir cette variable comme catégorique et de spécifier le nombre de niveaux à définition ou commande de programme. Le programme définit alors un vecteur de variables indicatrices. On fixe une variable indicatrice, disons x_{2dhij} , à 1 si l'observation hij appartient à la catégorie d , et on la fixe à 0 dans les autres cas. Pour éviter une dépendance linéaire entre les variables explicatives, le programme d'analyse reparamètre les indicateurs pour la variable catégorique. Cela se fait généralement par l'abandon de la catégorie finale de la variable catégorique; cette catégorie devient alors la catégorie de référence⁵. Le tableau XIX.2 illustre certains des effets qui peuvent être estimés pour le modèle 1 lorsque la variable dépendante est le revenu que le ménage tire de salaires, la variable explicative continue est le nombre de salariés du ménage et la variable catégorique définit le pays en quatre domaines régionaux (Nord, Sud, Est et Ouest).

45. Les coefficients de régression estimés pour les variables des domaines sont définis du point de vue de la différence entre un domaine et le domaine de référence. Le test de signification statistique d'un coefficient estimé pour le domaine Nord vise en fait à déterminer si le Nord et l'Ouest peuvent servir d'échantillons aléatoires de la même super-population commune. Si le coefficient correspondant à la région Nord s'écarte sensiblement de 0 (sur la base d'un test hypothétique de niveau de signification 0,05), l'analyste peut conclure qu'il est très improbable (5 % de chances ou moins) que les salaires des ménages du Nord et de l'Ouest puissent servir d'échantillons de la même super-population, une fois ajusté le nombre de salariés par ménage. Les programmes statistiques permettent aux utilisateurs de spécifier différents ensembles de références soit en classant les catégories (de telle façon que la catégorie de référence désirée soit la dernière) soit selon des spécifications explicites. Cela peut être un moyen utile d'obtenir des estimations significatives des paramètres de régression. D'autres comparaisons peuvent aussi être faites au moyen de fonctions des coefficients estimés.

⁵ On peut également estimer les coefficients des variables catégoriques en ajoutant une contrainte linéaire, par exemple en exigeant que la somme des effets ou la somme des effets pondérés soit nulle.

Tableau XIX.2
Interprétation des estimations des paramètres de régression linéaire
lorsque la variable dépendante correspond aux revenus salariaux du ménage
pour le modèle 1

Effet (généralement identifié dans les résultats du programme)	Coefficient de	Estimation de	Interprétation
Ordonnée à l'origine	$x_0 = 1$	α	Revenus salariaux du ménage aux niveaux de la cellule de référence ou zéro : 0 salarié dans la région Ouest
Salariés du ménage	x_{1hij}	β_1	Variation des revenus salariaux des ménages par salarié supplémentaire (ajustée en fonction de la région)
Région			Différences entre les revenus salariaux des ménages selon les régions (ajustées en fonction du nombre de salariés du ménage)
Nord ($d = 1$)	$X_{21hij} - X_{24hij}$	$\beta_2 = \gamma_1 - \gamma_4$	Nord par rapport à l'Ouest
Sud ($d = 2$)	$X_{22hij} - X_{24hij}$	$\beta_3 = \gamma_2 - \gamma_4$	Sud par rapport à l'Ouest
Est ($d = 3$)	$X_{23hij} - X_{24hij}$	$\beta_4 = \gamma_3 - \gamma_4$	Est par rapport à l'Ouest
Ouest (domaine de référence, $d = 4$)	$X_{24hij} - X_{24hij} = 0$	$\gamma_4 - \gamma_4 = 0$	Pas d'estimation

46. Le tableau XIX.3 montre des fonctions du modèle 1 qui peuvent être estimées sur la base des estimations des paramètres indiqués au tableau XIX.2. Le tableau XIX.3 présente des estimations des revenus salariaux des ménages du modèle 1 par région et par nombre de salariés par ménage. Ce nombre pourrait facilement atteindre trois salariés ou plus.

Tableau XIX.3.
Revenus salariaux des ménages qui peuvent être estimés (modèle 1)

Région	Par ménage comptant	
	Un salarié	Deux salariés
Nord	$\hat{\alpha} + \hat{\beta}_1 + \hat{\beta}_2$	$\hat{\alpha} + 2\hat{\beta}_1 + \hat{\beta}_2$
Sud	$\hat{\alpha} + \hat{\beta}_1 + \hat{\beta}_3$	$\hat{\alpha} + 2\hat{\beta}_1 + \hat{\beta}_3$
Est	$\hat{\alpha} + \hat{\beta}_1 + \hat{\beta}_4$	$\hat{\alpha} + 2\hat{\beta}_1 + \hat{\beta}_4$
Ouest	$\hat{\alpha} + \hat{\beta}_1$	$\hat{\alpha} + 2\hat{\beta}_1$

47. Examinons les suppositions que doit faire l'analyste lorsqu'il utilise le modèle 1 pour étudier les revenus salariaux des ménages. La supposition la plus critique est peut-être que les revenus salariaux des ménages sont liés de façon linéaire au nombre de salariés. La supposition de linéarité implique que les revenus salariaux du ménage augmentent du même montant lorsque le ménage passe de 0 à 1 salarié, de 1 salarié à 2, de 2 salariés à 3, etc. Cette supposition paraît douteuse. Comme les variables catégoriques nécessitent moins de suppositions quant au type de relation entre la variable explicative et la variable dépendante, l'analyste peut décider de convertir le nombre de salariés en variable catégorique et utiliser ainsi un modèle ne faisant intervenir que des variables catégorique⁶. L'équation ci-après est une variante du modèle 1 :

⁶ Pour un examen plus approfondi de la méthodologie d'évaluation de la bonne adaptation d'un modèle de régression linéaire et pour d'autres solutions de remplacement des relations non linéaires, prière de se reporter à Korn et Graubard (1999, pp. 95-100).

Modèle 2

$$y_{hij} = \alpha x_0 + \sum_{d=1}^{D_1} \gamma_{1d} x_{1dhij} + \sum_{d=1}^{D_2} \gamma_{2d} x_{2dhij} + \varepsilon_{hij}$$

48. Pour le modèle 2, l'analyste peut ne définir que deux catégories de salariés ou en définir un nombre beaucoup plus élevé selon la répartition du nombre de salariés dans les ménages. Pour limiter le nombre de paramètres à estimer, l'analyste peut se limiter à quatre catégories :

- Catégorie 1 : pas de salariés;
- Catégorie 2 : un salarié;
- Catégorie 3 : deux salariés;
- Catégorie 4 : trois salariés ou plus.

49. Ce modèle reste un modèle illustrant les principaux effets, mais le nombre de paramètres de régression est maintenant passé de cinq à sept. Le tableau XIX.4 illustre l'interprétation de l'estimation des coefficients de régression selon le modèle 2. Ce modèle ne nécessite plus que l'analyste suppose une relation linéaire entre les revenus salariaux et le nombre de salariés du ménage. Toutefois, comme le modèle ne comporte pas de termes d'interaction, il suppose que :

- L'effet des « salariés du ménage » est le même dans les quatre régions.
- « L'effet lié à la région » est le même quel que soit le nombre de « salariés du ménage ».

Tableau XIX.4

Interprétation des estimations des paramètres de régression linéaire

lorsque la variable dépendante indique les revenus salariaux des ménages, selon le modèle 2

Effet (généralement identifié dans les résultats du programme)	Coefficient de	Estimation de	Interprétation
Ordonnée à l'origine	$x_0 = 1$	α	Revenus salariaux du ménage aux niveaux de référence (0 salarié dans la région Ouest)
Salariés du ménage			
Un ($d = 1$)	$x_{11hij} - x_{14hij}$	$\beta_1 = \gamma_{11} - \gamma_{14}$	Un par rapport à deux
Deux ($d = 2$)	$x_{12hij} - x_{14hij}$	$\beta_2 = \gamma_{12} - \gamma_{14}$	Deux par rapport à aucun
Trois ou plus ($d = 3$)	$x_{13hij} - x_{14hij}$	$\beta_3 = \gamma_{13} - \gamma_{14}$	Trois par rapport à aucun
Aucun (domaine de référence, $d = 4$)	$x_{14hij} - x_{14hij} = 0$	$\gamma_{14} - \gamma_{14} = 0$	Pas d'estimation
Région			
Nord ($d = 1$)	$x_{21hij} - x_{24hij}$	$\beta_4 = \gamma_{21} - \gamma_{24}$	Nord par rapport à l'Ouest
Sud ($d = 2$)	$x_{22hij} - x_{24hij}$	$\beta_5 = \gamma_{22} - \gamma_{24}$	Sud par rapport à l'Ouest
Est ($d = 3$)	$x_{23hij} - x_{24hij}$	$\beta_6 = \gamma_{23} - \gamma_{24}$	Est par rapport à l'Ouest
Ouest (domaine de référence, $d = 4$)	$x_{24hij} - x_{24hij} = 0$	$\gamma_{24} - \gamma_{24} = 0$	Pas d'estimation

50. La plupart des modèles de régression permettent de tester les interactions entre les variables catégoriques. Dans ce cas, il y a neuf degrés de liberté pour l'interaction. Bien qu'il soit possible d'interpréter les effets des modèles de régression avec

deux principaux effets catégoriques et une interaction, nous recommandons une approche différente. Premièrement, déterminer s'il y a interaction : ici, on pourrait étendre le modèle 2 de manière à y inclure l'interaction entre « les salariés du ménage » et « la région ». Si le test statistique indique la présence d'interactions, incorporer le modèle complet avec 16 paramètres pouvant être estimés par l'application d'un modèle plus simple à une seule variable catégorique définie à 16 niveaux. Nous appellerons ce modèle 3 et l'exprimerons ainsi :

Modèle 3

$$y_{bij} = \alpha x_0 + \sum_{d=1}^{16} \beta_{1d} x_{1dbij} + \varepsilon_{bij}$$

51. Les 16 niveaux de la nouvelle variable catégorique et leurs estimations (entre parenthèses) sont les suivants :

- Nord, un salarié : $\hat{\alpha} + \hat{\beta}_1$
- Nord, deux salariés : $\hat{\alpha} + \hat{\beta}_2$
- Nord, trois salariés ou plus : $\hat{\alpha} + \hat{\beta}_3$
- Nord, pas de salariés : $\hat{\alpha} + \hat{\beta}_4$
- Sud, un salarié : $\hat{\alpha} + \hat{\beta}_5$
- Sud, deux salariés : $\hat{\alpha} + \hat{\beta}_6$
- Sud, trois salariés ou plus : $\hat{\alpha} + \hat{\beta}_7$
- Sud, pas de salariés : $\hat{\alpha} + \hat{\beta}_8$
- Est, un salarié : $\hat{\alpha} + \hat{\beta}_9$
- Est, deux salariés : $\hat{\alpha} + \hat{\beta}_{10}$
- Est, trois salariés ou plus : $\hat{\alpha} + \hat{\beta}_{11}$
- Est, pas de salariés : $\hat{\alpha} + \hat{\beta}_{12}$
- Ouest, un salarié : $\hat{\alpha} + \hat{\beta}_{13}$
- Ouest, deux salariés : $\hat{\alpha} + \hat{\beta}_{14}$
- Ouest, trois salariés ou plus : $\hat{\alpha} + \hat{\beta}_{15}$
- Ouest, pas de salariés : $\hat{\alpha}$

52. Avec la seizième catégorie définie comme la cellule de référence, l'estimation de l'ordonnée à l'origine du modèle 3 $\hat{\alpha}$ correspond à l'estimation des revenus salariaux de ménages pour cette cellule (ouest, pas de salariés). L'estimation des revenus salariaux pour chacune des 15 autres cellules est estimée comme l'estimation de la seizième cellule plus le coefficient de régression estimé pour cette cellule. On pourrait également obtenir ces 16 estimations à partir d'estimations directes. Si l'on applique les pondérations de l'enquête et la structure de conception à un logiciel d'enquête approprié, les estimations et leurs erreurs types estimatives devraient être identiques selon les deux approches (modèle 3 ou estimation directe). L'application du modèle 3 ne présente par d'avantages par rapport à l'établissement de 16 estimations directes.

53. Le choix d'échantillons de petite taille pour certaines des 16 cellules se fait au détriment de la précision des estimations pour ces cellules. En utilisant un modèle aux principaux effets (modèle 1 ou 2), on obtient des estimations plus précises pour les cellules utilisant de petits échantillons en « empruntant » la taille d'un échantillon aux estimations marginales et en faisant quelques suppositions supplémentaires (voir ci-dessus) sur la façon dont on détermine la population finie à partir de la super-population hypothétique.

54. Généralement, les analystes utilisent des modèles pour effectuer un ajustement en fonction d'un certain nombre de variables explicatives. Supposons qu'un

analyste souhaite effectuer un ajustement en fonction des caractéristiques d'une ville ou d'une communauté (pourcentage de population urbaine). Dans un modèle linéaire des principaux effets, l'ajustement en fonction du pourcentage de population urbaine (en tant que variable continue ou que variable explicative catégorique) donne une estimation des effets de la région, à supposer que le pourcentage de population urbaine (type) soit le même dans chaque région. Sans ajustement pour tenir compte des covariates, le modèle (ou les estimations directes) représente les paramètres régionaux tels qu'ils sont; avec un ajustement du modèle pour tenir compte des covariates, le modèle représente les paramètres régionaux tels qu'ils seraient si les effets des covariates étaient supprimés. Korn et Graubard (1999, pp. 126-140) examinent l'utilisation de marges de prédiction comme méthode de normalisation.

F. Modèles de régression logistique

55. La régression linéaire ne s'applique pas lorsque la variable dépendante est catégorique. Bien que nous disposions des procédures de modélisation multinomiales, nous n'examinerons que les variables catégoriques binaires (à deux niveaux) qui peuvent s'analyser à l'aide de modèles de régression logistique. En ce sens, la régression logistique est un cas spécial plus simple de régression multinominale.

56. Pour une variable à deux catégories ou binaire codée 0 ou 1, on peut appliquer la régression linéaire mais elle peut produire des valeurs différentes de 0 et 1. On peut utiliser la régression linéaire dans un premier temps avec une variable dépendante binaire pour identifier les variables explicatives qui sont de bons prédicteurs de la variable dépendante, surtout si les ensembles logiciels dont dispose l'analyste ont des procédures de sélection de variables incorporées au logiciel de régression linéaire mais pas au logiciel de régression logistique.

57. On utilise les méthodes numériques pour définir les modèles de régression logistique, de sorte qu'il peut parfois être difficile de parvenir à une solution. Il importe que les utilisateurs tiennent compte de tout avertissement donné par le logiciel lorsqu'il se pose des problèmes de convergence; en général, on peut résoudre ces problèmes en simplifiant le modèle.

58. On peut exprimer un modèle de régression logistique comportant une variable explicative continue et une variable explicative catégorique par l'équation suivante :

Modèle 4

$$\log \left(\frac{p(x_{hij})}{1 - p(x_{hij})} \right) = \alpha x_0 + \beta_1 x_{1hij} + \sum_{d=1}^D \gamma_d x_{2hij} + \varepsilon_{hij}$$

59. Hormis pour la variable dépendante, les termes du modèle 4 sont définis de la même manière que ceux du modèle 1. Pour comprendre la transformation logistique, considérons un exemple où $p(x_{hij})$ est fonction des variables explicatives; pour plus de commodité, on le désigne par p . Supposons en outre que p est la probabilité qu'un ménage auquel correspond un ensemble donné de valeurs pour les variables explicatives a un revenu qui se situe en dessous du niveau établi comme seuil de pauvreté. Dès lors, $p/(1 - p)$ correspond à la probabilité de se trouver en dessous de ce seuil, et $\log(p/1 - p)$ est le log-odds de p , parfois dénommé logit (p). Le modèle 4 vise à lier le log-odds de p aux x . Les observations portent sur des ménages où nous observons non pas la probabilité qu'ils se situent en dessous du seuil de pauvreté, mais leur

situation effective, indépendamment de ce seuil. Par ailleurs, comme la variable dépendante et un log-odds de p , chaque paramètre [α , β_1 et γ_d ($d = 1, \dots, D$)] figure aussi sur l'échelle des log-odds de p ; en outre, on suppose que la relation entre les log-odds de p et les x est linéaire (comparer au modèle 3 ci-dessus).

60. La reparamétrisation des variables explicatives catégoriques et la définition des catégories de référence sont les mêmes que pour la régression linéaire examinée ci-dessus. Les paramètres du modèle de régression des résultats du programme de régression logistique ressemblent à ceux de la régression linéaire, mais ils se prêtent à des interprétations différentes. Le tableau XIX.5 résume l'interprétation des estimations habituelles pour le modèle 4. On notera qu'il y a cinq paramètres estimatifs (une ordonnée à l'origine, α , et quatre β).

Tableau XIX.5

Interprétation des estimations des paramètres de régression logistique lorsque la variable dépendante est un indicateur des ménages se situant en dessous du seuil de pauvreté, avec le modèle 4

Effet (généralement identifié dans les résultats du programme)	Coefficient de	Estimation de	Interprétation
Ordonnée à l'origine t	$x_0 = 1$	α	Le log-odds de se situer en dessous du seuil de pauvreté à la cellule de référence ou aux niveaux zéro : 0 salarié dans la région Ouest
Salariés du ménage	x_{1hij}	β_1	Variation du log-odds de se situer en dessous du seuil de pauvreté par salarié supplémentaire (ajustée en fonction de la région)
Région			Différences régionales du log-odds de se situer en dessous du seuil de pauvreté (ajustées en fonction du nombre de salariés du ménage)
Nord ($d = 1$)	$x_{21hij} - x_{24hij}$	$\beta_2 = \gamma_1 - \gamma_4$	Nord par rapport à l'Ouest
Sud ($d = 2$)	$x_{22hij} - x_{24hij}$	$\beta_3 = \gamma_2 - \gamma_4$	Sud par rapport à l'Ouest
Est ($d = 3$)	$x_{23hij} - x_{24hij}$	$\beta_4 = \gamma_3 - \gamma_4$	Est par rapport à l'Ouest
Ouest (domaine de référence, $d = 4$)	$x_{24hij} - x_{24hij}$	$\gamma_4 - \gamma_4 = 0$	Pas d'estimation

61. On notera également que les paramètres du modèle logistique prédisent le log-odds de se situer en dessous du seuil de pauvreté mais ne prédisent pas directement la probabilité de se situer en dessous de ce seuil. Considérons β_2 au tableau XIX.5. Elle s'exprime comme suit, différemment du log-odds :

$$\beta_2 = \log\left(\frac{p(\text{Nord})}{1 - p(\text{Nord})}\right) - \log\left(\frac{p(\text{Ouest})}{1 - p(\text{Ouest})}\right)$$

Par les propriétés des logarithmes, elle peut aussi s'exprimer comme le log d'un odds ratio (rapport des cotes) :

$$\beta_2 = \log\left(\frac{\left(\frac{p(\text{Nord})}{1 - p(\text{Nord})}\right)}{\left(\frac{p(\text{Ouest})}{1 - p(\text{Ouest})}\right)}\right)$$

Le résultat type des procédures de régression logistique donne aussi les odds-ratios, qui peuvent se calculer directement comme :

$$e^{\beta_2} = \left(\frac{\left(\frac{p(\text{Nord})}{1-p(\text{Nord})} \right)}{\left(\frac{p(\text{Ouest})}{1-p(\text{Ouest})} \right)} \right)$$

En outre, le modèle permet de déterminer les probabilités que certains ménages se situent en dessous du seuil de pauvreté :

$$p(x_{hij}) = \frac{1}{1 + e^{-\log it [p(x_{hij})]}}$$

62. Lorsqu'ils se réfèrent aux résultats de l'ajustement du modèle (model fitting), les auteurs interprètent parfois un odds-ratio de 2 comme indiquant que la probabilité du fait (la pauvreté) dans un domaine (par exemple, le Nord) est deux fois celle de ce fait dans l'autre domaine (par exemple, l'Ouest). Si ce type d'interprétation est plus ou moins juste pour certains faits rares (p proche de 0), il l'est rarement pour des faits plus courants.

G. Utilisation de modèles à niveaux multiples

63. Nous examinerons à présent la question de la modélisation à niveaux multiples, et commencerons par souligner la nécessité de tenir compte de la structure des données de l'enquête. Il convient en effet de considérer la structure imposée par les enquêtes conçues pour être à phases multiples. Par exemple, les régions agroécologiques d'un pays peuvent former des strates à partir de chacune desquelles on peut choisir un certain nombre d'unités administratives. Ces dernières constituent alors les unités primaires d'échantillonnage. On choisit ensuite des unités secondaires à partir de chaque unité primaire, puis des unités du niveau suivant à partir des unités secondaires, et ainsi de suite. Cela produit une structure hiérarchique des données. Il se peut que des variables de stratification soient utilisées à un niveau ou plus.

64. Par exemple, une enquête portant sur les familles d'agriculteurs d'une région peut nécessiter l'utilisation des unités administratives de la région comme unités primaires, puis la sélection de certains villages à partir de chaque division, puis celle de ménages à partir de chaque village, en veillant éventuellement à inclure des ménages de différentes catégories de niveau de richesse. Ici, il faut tenir compte des différentes sources de variabilité des données recueillies au niveau des ménages. La variation globale comprend la variation entre les divisions administratives, la variation entre les villages et la variation entre les ménages d'un même village. Souvent, des données sont également recueillies à chaque niveau de la hiérarchie : ici, au niveau des ménages, au niveau des villages et au niveau des divisions administratives. Il est donc important d'identifier et de noter quelles variables sont mesurées au niveau des villages (par exemple, la présence d'un vulgarisateur ou l'octroi de subventions de l'État pour les engrais) et quelles variables sont mesurées au niveau des ménages (par exemple, les caractéristiques socio-économiques du ménage).

65. Aux fins de l'analyse des données, des fichiers sur tableurs « à plat » peuvent être établis pour recueillir les informations au niveau du village et au niveau du ménage, avec un identificateur pour lier les deux types de fichiers. Cette formule convient si les objectifs de l'analyse exigent que les données au niveau des villages soient ana-

lysées séparément de celles au niveau des ménages. Mais il est préférable de disposer d'une base de données relationnelle, c'est-à-dire d'une base de données qui permettent aux données de différents niveaux d'être mises en mémoire dans un fichier, avec des liens qui permettent d'établir une relation entre les données d'un niveau et celles d'un autre. L'analyse doit pouvoir tirer les informations des multiples niveaux afin d'étudier les relations entre les différents niveaux, de manière à permettre (par exemple) une interprétation d'ensemble.

66. Ici, la modélisation à niveaux multiples est la technique statistique clé. Cette approche de la modélisation (Goldstein, 2003; Snijders et Bosker, 1999; Kreft et De Leeuw, 1998) est souhaitable parce qu'elle permet d'explorer les relations à travers divers niveaux hiérarchiques et à l'intérieur d'un même niveau d'une conception à phases multiples, en tenant compte de la variabilité aux différents niveaux. Elle tient compte également des corrélations entre variables au même niveau. Elle permet aussi, par l'utilisation d'un logiciel approprié, par exemple *MLwiN* (Rasbash *et al.*, 2001) et SAS (2001), des erreurs types fondées sur le modèle pour les estimations d'enquêtes de conception complexe. Ces erreurs types peuvent offrir une approximation raisonnable d'erreurs types plus exactes qui tiennent compte de la stratification et de la mise en grappes. Il convient de noter que *MLwiN* peut également tenir compte des pondérations d'échantillonnage. C'est important, car des probabilités de sélection inégales dans un échantillonnage à phases multiples peuvent introduire une distorsion des estimateurs des principaux paramètres. Ces questions sont examinées de façon plus approfondie par Pfeiffermann *et al.* (1998) et Korn et Graubard (2003).

67. Il convient d'évoquer brièvement les conséquences qu'implique le fait de ne pas tenir compte de la structure hiérarchique, ce qui peut se produire lorsque les données sont agrégées à un niveau plus élevé ou désagrégées à un niveau inférieur. Si l'analyse est pertinente et n'est nécessaire qu'à un niveau, il n'y a pas de problème. En revanche, il faut bien veiller à ce que toutes inductions ne soient faites qu'à ce niveau. Il n'est pas possible de faire d'inductions au sujet d'un niveau particulier de la hiérarchie à partir de données analysées à un autre niveau. Autrement dit, une analyse ne faisant pas cas de la hiérarchie ne permet pas d'explorer les effets à travers les divers niveaux. Il se présente une autre difficulté si l'on analyse les données au niveau le plus bas en considérant les unités du niveau plus élevé comme un facteur dans l'analyse. C'est inefficace car on ne peut généraliser les conclusions pour toutes les unités des niveaux plus élevés de la population : ces conclusions ne s'appliquent qu'aux unités sélectionnées.

68. Nous présentons ci-dessous un scénario illustrant la façon dont l'utilisation de la modélisation à niveaux multiples peut être utile pour l'analyse des relations. D'autres exemples sont données par Congdon (1998), Langford, Bentham et McDonald (1998) et Goldstein *et al.* (1993).

Exemple 1

69. Dans une étude des facteurs contribuant au succès de la cogestion communautaire des ressources côtières des pays insulaires du Pacifique, 31 sites ont été choisis dans cinq pays et 133 entrevues ont été effectuées avec de petits groupes de réflexion composés de deux à six répondants appartenant à différents ménages (Banque mondiale, 2000). Fidji, Palau, Samoa, les îles Salomon et Tonga étaient les pays choisis pour représenter tout un ensemble de conditions de gestion des ressources côtières. L'étude a permis de recueillir des « perceptions de succès » fondées sur les tendances perçues de la prise par unité d'effort [*perceived catch per unit effort* (CPUE)], les conditions d'habitat, les menaces pesant sur le site et une évaluation du respect des règles. Les

trois premiers indicateurs ont été mesurés sur une échelle à cinq points (5 = amélioration sensible; 1 = dégradation sensible) et le respect des règles a été mesuré sur une échelle à quatre points.

70. On a également recueilli des données à l'échelon national auprès des ministères de la pêche et de l'environnement de chaque pays, et au niveau des sites. Par ailleurs, chaque groupe de réflexion, composé de membres de plusieurs ménages, a été invité à donner ses perceptions au sujet d'un ensemble de trois ressources (CPUE), de trois types d'habitat, de trois menaces et de cinq règles de gestion à respecter. Les informations ainsi recueillies au cours de l'étude se situaient à quatre niveaux : le pays, le site, le groupe de réflexion et certains types donnés de ressources, d'habitat, de menace ou de règle.

71. Il importe toutefois de noter que cette enquête a utilisé un échantillonnage non probabiliste; on peut donc arguer que toutes conclusions analytiques ne peuvent être généralisées à aucune population cible clairement définie. Toutefois, pour les besoins de la présente analyse, nous supposons que l'échantillonnage a été effectué sur une base probabiliste et que le groupe de réflexion analyserait les données à l'aide d'un modèle à multiples niveaux; la variable particulière étudiée étant la perception de la tendance concernant la CPUE, obtenue en prenant la moyenne des niveaux de perception pour les trois ressources. L'effet lié au pays (au niveau le plus élevé de la hiérarchie) pourrait être inclus dans le modèle comme facteur (effet fixe) car il s'agit essentiellement d'une variable de stratification. Toutefois, pour permettre de généraliser les résultats pour tous les sites cogérés, il faut inclure les sites comme variable aléatoire plutôt que comme effet fixe. Les groupes de réflexion au niveau des sites entreraient également le modèle comme effet aléatoire. L'essence de la modélisation à niveaux multiples réside dans l'inclusion d'un ensemble de variables à effet fixe et de variables à effet aléatoire. Ces modèles permettent également d'étudier les interactions entre les variables au niveau des sites et les variables au niveau du groupe de réflexion.

72. Pour illustrer la façon dont on peut formuler le modèle à niveaux multiples pour répondre à certaines questions spécifiques de l'enquête, nous utilisons un exemple tiré d'une Enquête sur la production et la sécurité alimentaires effectuée en 2000-2001 au Malawi. Cette enquête visait à évaluer un programme qui avait pour but d'accroître la sécurité alimentaire parmi les petits exploitants ruraux par la distribution d'un lot de démarrage comprenant de l'engrais et des semences de maïs et de légumineuses.

Exemple 2

73. L'Enquête sur la production et la sécurité alimentaires était une enquête nationale qui a utilisé un échantillonnage stratifié à deux phases en prenant les districts comme strates. Quatre villages ont été sélectionnés dans chacun des 27 districts du Malawi, et une trentaine de ménages ont été sélectionnés dans chaque village. La sélection des villages a été limitée à ceux qui comptaient plus de 40 ménages (afin qu'il y ait assez de ménages dans le village pour permettre d'interroger les bénéficiaires du lot de démarrage) et moins de 250 ménages (pour que l'équipe puisse faire son travail dans le délai imparti et avec les ressources disponibles)⁷. Une fois cette restriction admise, l'échantillonnage à chaque niveau s'est fait de façon aléatoire. Au total, 108 villages ont été visités et 3 030 ménages ont été interrogés dans le cadre de l'enquête.

74. Les données que nous considérons pour la modélisation à niveaux multiples proviennent d'un questionnaire rempli par les ménages pour les besoins de l'enquête. Le sous-ensemble de variables que nous examinerons dans notre illustration sont le district, le village, le numéro d'identification du ménage, le sexe et l'âge du chef

⁷ Cette contrainte sur la population cible limite l'inférence à la population résidant dans les villages se situant dans ces limites de taille.

de famille, la taille de la famille, la question de savoir si le ménage a ou non reçu le lot de démarrage et deux indices reflétant les biens⁸ et le revenu du ménage⁹.

75. On peut utiliser plusieurs modèles à niveaux multiples avec ces données. Quand on conçoit le modèle, la première chose est de décider quelles sont les variables aléatoires et quelles sont celles aux effets fixes.

76. Dans l'exemple 2, le district est une variable de stratification et doit être considéré comme un effet fixe. En général, tout effet est considéré comme fixe si des répétitions du processus d'échantillonnage donne le même ensemble de sélections. Par ailleurs, les villages et les ménages ayant été sélectionnés au hasard, ils ont des effets aléatoires sur le modèle.

77. Le modèle de base pour analyser (disons) l'indice de biens (AI) est :

Modèle 5

$$y_{ijk} = \mu + d_k + U_{jk} + \varepsilon_{ijk}$$

où d_k est l'effet lié au district ($k = 1, 2, \dots, 27$), et les indices i et j correspondent respectivement au i^{e} ménage et au j^{e} village. Il est parfois utile de considérer le paramètre du district comme reflétant l'écart de la valeur moyenne de AI pour le district k de la valeur moyenne globale de AI pour tous les districts. Toutefois, le logiciel de modélisation utilise une paramétrisation différente et fixe l'un des effets liés au district à zéro. Les autres effets permettent alors des comparaisons entre les valeurs AI de chaque district avec la valeur AI du district dont l'effet a été fixé à zéro.

78. Dans ce modèle, U_{jk} et ε_{ijk} sont des variables aléatoires qui représentent respectivement la variation entre tous les villages du district k (dont on suppose qu'elle est la même pour tous les districts) et la variation entre tous les ménages du village j du district k (dont on suppose qu'elle est la même pour toutes les combinaisons de villages et de districts). U_{jk} et ε_{ijk} sont des variables aléatoires dont on suppose dans le modèle qu'elles sont normalement distribuées avec une moyenne zéro et des variances constantes σ_u^2 et σ_e^2 , respectivement. On suppose en outre qu'elles sont indépendantes les unes des autres. Nous pouvons donc écrire : $U_{jk} \sim N(0, \sigma_u^2)$ et $\varepsilon_{ijk} \sim N(0, \sigma_e^2)$.

79. L'utilisation de ce modèle permet d'établir des estimations de σ_u^2 et σ_e^2 et des estimations de d_k , ainsi que des erreurs types pertinentes. Les estimations des paramètres pour d_k ($k=1, 2, \dots, 27$), permettent de comparer les moyennes de AI à travers tous les districts.

80. Supposons à présent que nous cherchions à déterminer comment la taille du ménage influe sur la variation de AI (variable quantitative) et si le ménage a ou non reçu un lot de démarrage (variable binaire). Cela serait inclus dans le modèle de la même manière qu'avec une modélisation linéaire générale type. Le modèle serait donné par l'équation :

Modèle 6

$$y_{ijk} = \mu + d_k + U_{jk} + t_{p(ijk)} + \beta x_{ijk} + \varepsilon_{ijk}$$

où $t_{p(ijk)}$ représente l'effet correspondant à la réception du lot de démarrage; x_{ijk} représente la taille du ménage et β représente la pente décrivant la relation de x_{ijk} avec y_{ijk} , c'est-à-dire la relation entre la taille du ménage (HHSIZE) et l'indice de biens (AI).

81. Ici, $t_{p(ijk)}$ et β sont tous deux considérés comme des effets fixes, tandis que U_{jk} et ε_{ijk} sont (comme auparavant) des effets aléatoires. La forme de ce modèle suppose que la relation entre HHSIZE et AI est la même dans tous les villages et tous les districts.

⁸ L'indice des biens est une moyenne pondérée de différents nombres de têtes de bétail et de différents biens du ménage; par exemple, radio, bicyclette, charrette à bœufs, etc.

⁹ L'indice de revenu est fondé sur les revenus provenant de différentes sources.

82. L'inclusion de ces deux composantes de variation (village et ménage) dans le modèle ci-dessus signifie que le modèle tient compte de la variabilité aux deux niveaux de la hiérarchie. Cela veut dire que les erreurs types liées à $t_{p(ijk)}$ et β sont calculées correctement, comme le seraient les résultats de tests de signification liés à ces paramètres. Tel n'aurait pas été le cas si l'on avait utilisé un modèle général linéaire considérant les villages comme effets fixes. Même si l'on utilisait un logiciel d'enquête (tenant compte des pondérations d'échantillonnage), les procédures de régression linéaire ne tiendraient pas compte de la corrélation entre les ménages d'un même village.

83. Il y a un autre avantage important à traiter les villages comme effets aléatoires. Si les villages avaient été considérés comme fixes, les conclusions de l'analyse ne se seraient appliquées qu'à l'ensemble de villages visités dans le cadre de l'enquête. Le fait de considérer les villages comme effets aléatoires signifie que les conclusions concernant la relation entre la taille des ménages et le AI, la comparaisons des AI entre les ménages ayant ou non reçu le lot de démarrage et la comparaison entre districts peuvent toutes être généralisées à l'ensemble de la population de tous les villages. On peut aussi étudier l'interaction entre l'effet fixe au niveau du district d_k et l'effet lié au fait de recevoir le lot de démarrage $t_{p(ijk)}$ en incluant ce terme d'interaction dans le modèle.

84. On obtient un autre modèle utile en considérant le terme de pente β du modèle 6 comme une variable aléatoire qui varie d'un village à l'autre. Ce modèle est souvent appelé modèle de régression du coefficient aléatoire. Il s'exprime alors par l'équation :

Modèle 7

$$y_{ijk} = \mu + d_k + U_{jk} + t_{p(ijk)} + \beta_j x_{ijk} + \varepsilon_{ijk}$$

où β_j est supposé $N(\beta, \sigma_\beta^2)$. En outre, comme β_j est aléatoire entre tous les villages, on peut également le considérer comme ayant une covariance avec U_{jk} , disons $\sigma_{\beta u}$.

85. Ainsi, dans l'analyse présentée ici, le fait de tester l'hypothèse selon laquelle σ_β^2 est nul nous indique effectivement s'il y a variabilité de la pente de la relation AI-HHSIZE entre les villages. Si cette hypothèse ne peut être rejetée, on peut conclure que la forme de la relation est la même pour tous les villages.

86. On peut étendre ce modèle aux variables au niveau des villages; par exemple, l'accès à une alimentation en eau potable ou le degré d'accès aux conseils d'agents de vulgarisation agricole. Ici, les avantages réels de la modélisation à niveaux multiples entrent en jeu car il serait alors possible d'explorer les relations entre ces variables au niveau des villages et les variables au niveau des ménages. Autrement dit, la modélisation à niveaux multiples rend possible l'étude des relations entre variables de différents niveaux de l'échantillonnage hiérarchique. Cela offre pour avantages de permettre de tenir compte de la structure de corrélation entre les unités de niveau inférieur lorsque l'on analyse conjointement les variables aux différents niveaux. Dans l'exemple ci-dessus, on pourrait envisager d'autres modèles, par exemple, des modèles tenant compte du sexe et de l'âge du chef de famille et des interactions entre ces termes et les termes qui figuraient déjà dans le modèle.

87. Bien entendu, l'utilisation de modèles à niveaux multiples comporte des limitations. De même qu'avec les autres procédures de modélisation, le modèle à niveaux multiple imaginé ici est censé être raisonnablement correct et conforme à la conception de l'échantillon. La question de savoir si ces hypothèses sont justes peut certes se discuter.

H. Modélisation à l'appui des processus d'enquête

88. Même lorsque l'enquête sur les ménages est utilisée strictement pour donner des statistiques descriptives, la modélisation peut s'avérer nécessaire à l'appui d'autres processus d'enquête. Les ajustements pour tenir compte des cas de non-réponse sont souvent fondés directement ou indirectement sur des modèles statistiques : Groves *et al.* (2002, pp. 197-443) examinent divers moyens de prendre en compte la non-réponse, qui supposent tous le recours à un certain modèle statistique. Les modèles de régression logistique peuvent être utilisés pour prédire l'inclination à répondre et effectuer les ajustements nécessaires ou pour déterminer les catégories de pondérations en fonction de cette inclination [voir, par exemple, Folsom (1991); Folsom et Witt (1994); ou Folsom et Singh (2000)]. On peut également utiliser les modèles statistiques prédictifs dans le cadre de la procédure d'imputation des données manquantes [voir, par exemple, Singh, Grau et Folsom (2002)]. Enfin, les modèles statistiques peuvent servir à évaluer les expériences méthodologiques effectuées dans le cadre des enquêtes [voir, par exemple, Hughes *et al.* (2002)].

I. Conclusions

89. Dans ce chapitre, nous avons cherché à analyser les questions qui découlent de l'analyse des données d'enquêtes. Parmi ces questions figure l'utilisation des pondérations et de méthodes appropriées d'estimation de variance avec les approches descriptives et analytiques des données d'enquêtes. Nous donnons également un aperçu des situations pratiques où les techniques de modélisation ont un rôle à jouer dans l'analyse de ces données. Ces techniques sont des outils précieux mais leur utilisation exige que l'on accorde une grande attention aux hypothèses sous-jacentes.

90. Nous avons examiné le rôle des pondérations et l'importance de la structure de l'échantillon dans l'établissement de statistiques descriptives et analytiques des données d'enquêtes. Les logiciels d'analyse de ces données, qui utilisent des pondérations et tiennent compte de la structure de l'échantillon, peuvent être utilisés pour estimer les paramètres des modèles de régression linéaire et logistique sur la base de ces données. Les estimations fondées sur l'échantillon sont des estimations de ce qui serait obtenu si l'on appliquait ces modèles à l'ensemble de la population finie. Par ailleurs, on peut aussi déterminer les erreurs types de ces estimations. Les variables explicatives des modèles de régression appliqués aux données de l'enquête sont presque toujours observées telles qu'elles se présentent dans la population plutôt qu'attribuées de façon aléatoire selon une certaine conception expérimentale. Les analystes doivent bien comprendre que les coefficients de régression fondés sur les données d'enquêtes ne reflètent que les relations qui existent entre la variable dépendante et les variables explicatives de la population et n'impliquent pas nécessairement une relation de cause à effet. Nous avons examiné comment les paramètres de régression et les modèles de régression logistique se rapportent à de simples statistiques descriptives et comment ils peuvent s'interpréter pour certains modèles relativement simples.

91. La modélisation à niveaux multiples, en particulier, est généralement considérée comme une technique assez « avancée » utilisée de préférence en consultation avec un statisticien rompu à son utilisation et conscient des limites de cette technique. Actuellement, les modèles à niveaux multiples semblent rarement utilisés pour analyser les enquêtes effectuées dans les pays en développement; cependant, leur utilisation serait particulièrement souhaitable pour leur pénétration des relations entre variables de différents niveaux et leur aptitude à tenir compte de la variabilité entre les unités d'échantillonnage aux différents niveaux d'une conception à phases multiples.

92. Nous avons montré qu'il n'est pas trop difficile de formuler des modèles à niveaux multiples pour quiconque est rompu à l'utilisation de modèles linéaires généraux [*general linear models* (GLM)]; mais là encore, il faut vérifier certaines hypothèses liées aux modèles en effectuant certaines analyses, comme avec les GLM. La modélisation à niveaux multiples est également possible lorsque la principale réponse est binaire, mais nous n'avons pas présenté d'exemple d'un tel cas. Il faut également faire preuve de discernement pour déterminer quels effets sont aléatoires et quels effets sont fixes, et comment la définition du modèle doit permettre de répondre à certains objectifs précis de l'enquête.

93. Néanmoins, de même qu'avec toutes les techniques statistiques, les méthodes de modélisation examinées dans ce chapitre présentent diverses limitations dont il faut tenir compte quand on les utilise. Nous avons recommandé d'utiliser des pondérations et un logiciel d'analyse qui tiennent compte de la structure de l'échantillon. Or, il faut savoir qu'il est difficile de trouver le logiciel approprié à cet égard. Le chapitre XXI décrit plusieurs logiciels qui tiennent compte des questions de conception de l'échantillon du point de vue des procédures de régression multiple et de régression logistique. Malheureusement, ces logiciels ne sont pas adaptés aux modèles à niveaux multiples. Il est donc recommandé à l'utilisateur d'opter pour un logiciel statistique d'usage plus général, tel que SAS (2001), GenStat (2002) ou SPSS (2001), ou pour un logiciel spécialisé tel que *MLwiN* (Rasbash *et al.*, 2001).

94. Dans ce chapitre, nous avons proposé quelques techniques de modélisation qui peuvent être utiles pour l'analyse des données d'enquêtes. Nous recommandons aux analystes et aux chercheurs d'envisager sérieusement d'utiliser ces méthodes, lorsqu'elles conviennent aux objectifs de l'enquête, lors de l'analyse des données d'enquêtes, pour extraire le plus d'informations possible des données d'enquêtes recueillies à grands frais.

REMERCIEMENTS

L'auteur tient à adresser des remerciements tout particuliers aux réviseurs et éditeurs pour leurs commentaires et suggestions utiles, et plus particulièrement à M. Graham Kalton pour ses contributions au débat sur la pondération.

Il remercie également le Department for International Development (DFID) du Gouvernement du Royaume-Uni de Grande-Bretagne et d'Irlande du Nord d'avoir permis de recueillir des idées pour ce chapitre en finançant de nombreux projets intéressants comportant des enquêtes effectuées dans des pays en développement. Néanmoins, les informations contenues dans ce chapitre n'engagent que la responsabilité de ses auteurs et n'impliquent en rien l'expression d'une opinion quelle qu'elle soit de la part du DFID.

RÉFÉRENCES

- Banque mondiale (2000). *Voices from the Village: A Comparative Study of Coastal Management in the Pacific Islands*. Pacific Islands Discussion Paper Series, No. 9. Washington, D.C., Banque mondiale. Papua New Guinea and Pacific Islands Country Management Unit.
- Chambers, R. L., et C. J. Skinner (2003). *Analysis of Survey Data*. Chichester, Royaume-Uni, Wiley.

- Chromy, James R. (1998). *The Effects of Finite Sampling on State Assessment Sample Requirements*. Palo Alto, Californie, NAEP Validity Studies, American Institutes for Research.
- Cochran, W. G. (1977). *Sampling Techniques*, 3^e éd. New York, John Wiley and Sons.
- Congdon, P. (1998). A multi-level model for infant health outcomes: maternal risk factors and geographic variation. *The Statistician*, vol. 47, partie 1, pp. 159-182.
- Deville, J. C., et C. E. Särndal (1992). Calibration estimating in survey sampling. *Journal of the American Statistical Association*, vol. 87; pp. 87 et 376-382.
- Folsom, Ralph E., Jr. (1991). Exponential and logistic weight adjustments for sampling and non-response error reduction. Dans *Proceedings of the Social Statistics Section, American Statistical Association*. Alexandria, Virginie, American Statistical Association, pp. 376-382.
- _____, et Michael B. Witt (1994). Testing a new attrition non-response adjustment method for SIPP. Dans *Proceedings of the Survey Research Methods Section, American Statistical Association*, pp. 428-433.
- _____, et A. C. Singh (2000). The general exponential model for sampling weight calibration for extreme values, non-response and post-stratification. Dans *Proceedings of the Survey Research Methods Section, American Statistical Association*. Indianapolis, Indiana.
- GenStat (2002). *GenStat for Windows*, 6^e éd. Oxford, Royaume-Uni, VSN International, Ltd.
- Goldstein, H. (2003). *Multi-level Statistical Models*, 3^e éd. Londres, Arnold.
- _____, et al. (1993). A multi-level analysis of school examination results. *Oxford Review of Education*, vol. 19, pp. 425-433.
- Graubard, B. I., et E. L. Korn (2002). Inferences for super-population... *Statistical Science*, vol. 17, pp. 73-96.
- Groves, Robert M., et al. (2002). *Survey non-response*. New York, John Wiley and Sons, Inc.
- Horvitz, D. G., et D. J. Thompson (1952). A generalization of sampling without replacement from a finite universe. Dans *Journal of the American Statistical Association*, vol. 47, pp. 663-685.
- Hughes, Arthur, et al. (2002). Impact of interviewer experience on respondent reports of substance use. Dans *Redesigning an Ongoing National Household Survey: Methodological Issues*, J. Gfoerer, J. Eyerman and J. Chromy, eds. DHHS Publication, No. SMA 03-3768. Rockville, Maryland, Substance Abuse and Mental Health Services Administration, Office of Applied Studies, pp. 161-184.
- Kish, Leslie (1965). *Survey Sampling*. New York, John Wiley and Sons, Inc.
- Korn, E. L., et B. I. Graubard (1999). *Analysis of Health Surveys*. New York, John Wiley and Sons.
- _____, (2003). Estimating variance components by using survey data. Dans *Journal of the Royal Statistical Society B*, vol. 66, pp. 175-190.
- Kreft, I., et J. de Leeuw (1998). *Introducing Multi-level Modeling*. Londres.
- Langford, I. H., G. Bentham et A. McDonald (1998). Multi-level modeling of geographically aggregated health data: a case study on malignant melanoma mortality and UV exposure in the European community. *Statistics in Medicine*, vol. 17, pp. 41-58.
- Levy, S., et C. I. Barahona (2001). The targeted inputs programme, 2000-01: Rapport principal, non publié.
- Pfeffermann, D., et al. (1998). Weighting for unequal selection probabilities in multi-level models. In *Journal of the Royal Statistical Society B*, vol. 60, pp. 23-40.

- Rasbash, J., *et al.* (2001). *MLwiN Version 1.10.0007*. Multi-level Models Project. Londres, Institute of Education, University of London.
- Research Triangle Institute (2002). *SUDAAN User's Manual, Release 8.0*. Research Triangle Park, Caroline du Nord, Research Triangle Institute.
- Rubin, Donald B. (1987). *Multiple Imputation for Non-response in Surveys*. New York, John Wiley and Sons.
- SAS (2001). *SAS Release 8.2*. Cary, Caroline du Nord, SAS Institute, Inc., SAS Publishing.
- Singh, Avinash, Eric Grau et Ralph Folsom Jr. (2002). Predictive mean neighborhood imputation for NHSDA substance use data. Dans *Redesigning an Ongoing National Household Survey: Methodological Issues*, J. Gfroerer, J. Eyerman and J. Chromy, eds. DHHS publication, No. SMA 03-3768. Substance Abuse and Mental Health Services Administration, Office of Applied Studies.
- Skinner, C. J., D. Holt et T. M. F. Smith, eds. (1989). *Analysis of Complex Surveys*. New York, Wiley.
- Snijders, T. A. B., et R. J. Bosker (1999). *Multi-level Analysis: An Introduction to Basic and Advanced Multi-level Modelling*. Londres, Sage.
- SPSS (2001). *SPSS for Windows. Release 11.0*. Chicago, Illinois, LEAD Technologies, Inc.
- Tukey, J. W. (1977). *Exploratory Data Analysis*. Reading, Massachusetts, Addison-Wesley.
- Woodruff, R. S. (1971). A simple method for approximating the variance of a complicated sample. Dans *Journal of the American Statistical Association*, vol. 66, pp. 411-414.

Chapitre XX

Approches plus avancées de l'analyse des données d'enquêtes

GAD NATHAN

Université hébraïque
Jérusalem, Israël

RÉSUMÉ

Dans le présent chapitre, nous examinons les effets de l'utilisation pratique d'un échantillon de conception complexe dans la plupart des enquêtes par sondage sur l'analyse des données d'enquêtes. Nous indiquons les cas où la conception peut ou non influencer sur l'analyse et nous définissons les concepts essentiels. Une fois qu'un modèle a été choisi pour l'analyse, nous examinons les relations possibles entre le modèle et la conception de l'échantillon. Lorsque la conception peut avoir un effet sur l'analyse et que l'on ne peut ajouter une variable explicative supplémentaire liée à la conception, on peut utiliser essentiellement deux méthodologies : l'analyse classique, qui ne peut être modifiée en fonction de la conception; ou un nouvel outil analytique, qui peut être mis au point pour chaque type de conception. Différentes approches sont illustrées par l'application de données réelles à la régression linéaire, aux modèles linéaires et à l'analyse des données catégoriques.

Termes clés : échantillon de conception complexe, analyse des données d'enquêtes, régression linéaire, modèles linéaires, analyse des données catégoriques, analyse fondée sur un modèle.

A. Introduction

1. Conception de l'échantillon et analyse des données

1. La grande majorité des enquêtes par sondage, tant dans les pays développés que dans les pays en développement, ont essentiellement un but descriptif, à savoir donner des estimations de points et d'intervalles de mesures descriptives d'une population finie, telles que moyennes, valeurs médianes, fréquence de distribution et tabulation croisée de variables qualitatives. Néanmoins, comme l'ont montré les chapitres XV-XIX et comme le montrera le chapitre XXI, on constate une volonté croissante de tirer des conclusions au sujet des relations entre les variables étudiées plutôt que de décrire simplement certains phénomènes.

2. Dans le présent chapitre, nous nous efforcerons d'évaluer les effets d'échantillons de conception complexe couramment utilisés sur l'analyse des données d'enquêtes. Nous essaierons d'identifier des cas où la conception peut influencer sur l'analyse. En général, la conception de l'échantillon est sans effet sur l'analyse lorsque les variables sur lesquelles repose cette conception sont incluses dans le modèle analytique.

Souvent, toutefois, certaines de ces variables ne figurent pas dans le modèle, soit en raison d'un manque de spécifications soit du fait du manque d'intérêt porté à ces variables en tant que facteurs explicatifs. Cela peut entraîner de sérieuses distorsions.

3. Nous examinerons deux méthodologies essentielles concernant la façon de traiter les données d'un échantillon complexe lorsque certaines variables additionnelles liées à la conception ne sont pas incluses dans l'analyse. La première modifie un instrument analytique classique conçu pour traiter les données d'un échantillon aléatoire simple. La seconde présente un nouvel outil analytique spécifiquement conçu pour les échantillons de conception complexe.

4. Dans les paragraphes qui suivent, nous présentons quelques exemples d'effets possibles de la conception de l'échantillon sur l'analyse, puis nous définissons quelques concepts essentiels et nous discutons le rôle des effets de conception dans l'analyse des données d'échantillons complexes. La section B décrit les deux approches fondamentales de l'analyse de données d'échantillons complexes. Dans les sections C et D, nous examinons des exemples se rapportant respectivement à l'analyse de données continues et de données catégoriques. La section finale contient un résumé et quelques conclusions. Des définitions formelles et des résultats techniques sont présentés à l'annexe à ce chapitre.

2. Exemples des effets (et de l'absence d'effets) de la conception de l'échantillon sur l'analyse

5. Pour montrer les effets potentiels de la conception de l'échantillon sur l'analyse, nous examinerons l'exemple suivant, simple mais édifiant (pour plus de détails, voir Nathan et Smith, 1989). Soit Y la variable étudiée et X une variable auxiliaire. Supposons que le modèle de régression linéaire

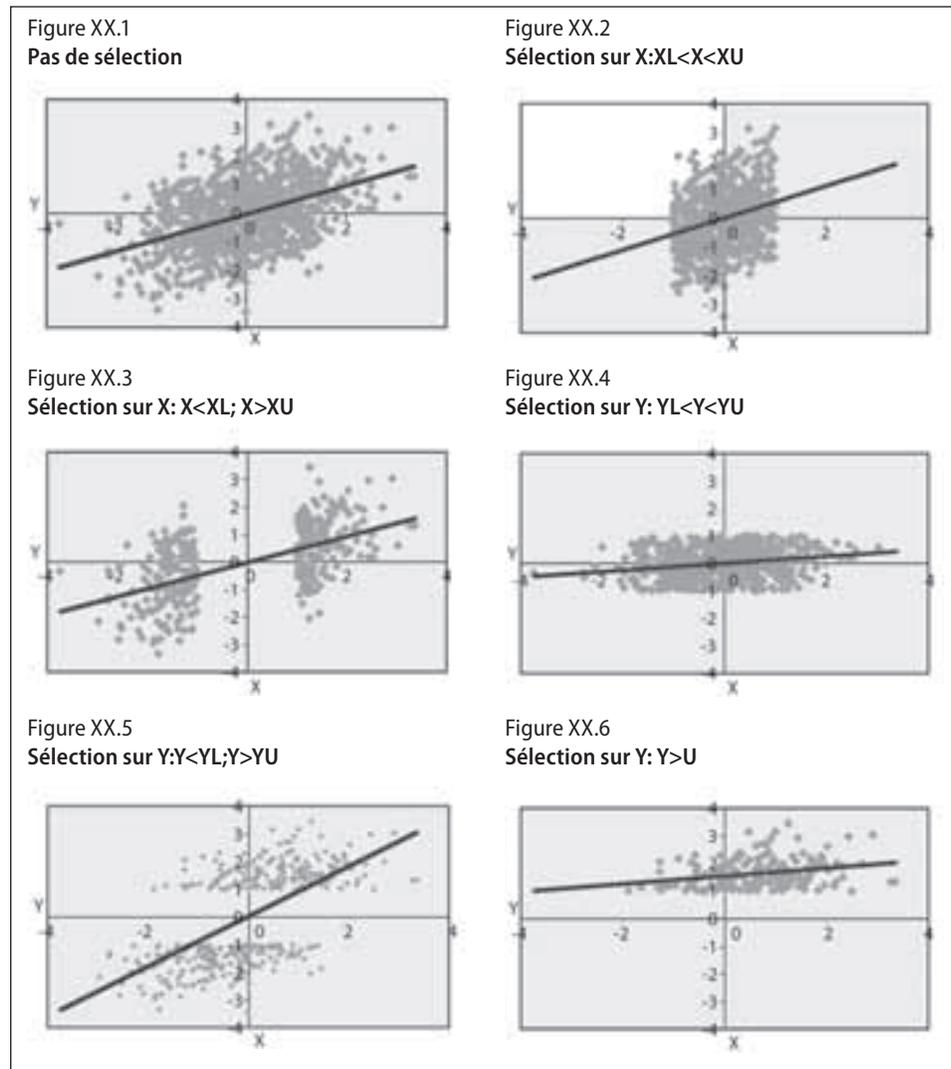
$$Y_i = \alpha + \beta X_i + \varepsilon_i, \text{ avec } \varepsilon_i | X_i \underset{\text{ind}}{\sim} N(0, \sigma_\varepsilon^2),$$

s'applique à la population. Ce modèle convient également pour tout échantillon aléatoire simple prélevé sur la population. Parfois, la supposition d'indépendance pour l'élément $\varepsilon_i | X_i$ convient mieux pour l'échantillon aléatoire simple qu'à la population d'où il est tiré. Par exemple, dans une population, les valeurs de ε_i peuvent être corrélées pour différents membres de la même famille, tandis que dans un échantillon aléatoire simple d'individus, où la probabilité que plus qu'un membre soit sélectionné par famille est très faible, la corrélation serait négligeable.

6. Avec un échantillonnage aléatoire simple, l'estimation type du coefficient de régression n'est pas entaché de distorsion. Le graphique de la figure XX.1 décrit Y par rapport à X pour la population totale; la figure serait la même pour un échantillon aléatoire simple. Pour les cinq figures XX.2 à XX.6, les échantillons sont choisis parmi la population selon des méthodes très différentes de celle de l'échantillonnage aléatoire simple. Considérons la sélection de l'échantillon uniquement sur la base de la valeur de X , par exemple en supprimant les points correspondant aux valeurs de X au-delà (ou à l'intérieur) de limites fixées, comme aux figures XX.2 et XX.3. Ces figures montrent clairement que cette sélection est sans effet sur l'estimation des paramètres de régression de l'ordonnée à l'origine (α) et de la pente (β), mais qu'elle peut influencer sur la variance des estimateurs.

7. Considérons à présent la sélection d'un échantillon sur la base des valeurs de la variable cible, Y , par exemple, en supprimant les points correspondant aux valeurs de Y au-delà (ou à l'intérieur) de limites fixées, comme aux figures XX.4-XX.6. Dans ces cas, il est clair que les estimations des pentes de la régression se trouvent faussées. Dans le dernier cas (figure XX.6), la suppression n'est pas symétrique, et l'estimation

de l'ordonnée à l'origine est également faussée. Ces exemples sont extrêmes car la sélection fondée sur la suppression de la variable dépendante est rare dans les enquêtes par sondage. En revanche, elle est très courante dans les études expérimentales ou d'observation, telles que les études de contrôle des cas en épidémiologie ou les études fondées sur un choix en économie [voir, par exemple, Scott et Wild (1986) et Manski et Lerman (1977)]. Néanmoins, dans de nombreux cas, les échantillons sont sélectionnés sur la base de variables qui peuvent être étroitement liées à la variable dépendante. C'est pourquoi une procédure d'échantillonnage courante, largement utilisée dans les enquêtes sur les établissements et les exploitations agricoles, consiste à sélectionner les unités selon une probabilité proportionnelle à leur taille. La mesure de la taille, par exemple, la production de l'année précédente, est évidemment liée à la variable étudiée lorsque celle-ci est la production de l'année en cours. Les estimations types des paramètres du modèle, tels que les coefficients de régression, peuvent être faussés lorsque l'on ne tient pas compte de la conception de l'échantillon.



8. Les exemples ci-dessus illustrent les dangers que comporte le fait d'effectuer une analyse fondée sur les données d'un échantillon complexe en faisant comme si ces données provenaient d'un échantillonnage aléatoire simple. Ces exemples montrent

qu'il est nécessaire de déterminer quand il est probable que la conception influe sur l'analyse et de tenir compte de cette conception lorsque tel est le cas.

3. Concepts de base

9. La plupart des enquêtes par sondage sont conçues principalement à des fins descriptives (ou aux fins d'énumération). Elles visent à estimer les valeurs des paramètres d'une population finie, tels que le revenu médian des ménages ou la proportion d'adultes ayant le sida. Ce sont là des statistiques qui, en principe, peuvent se mesurer avec exactitude si l'on englobe toute la population dans l'enquête, c'est-à-dire s'il s'agit d'un recensement de la population plutôt que d'un échantillon aléatoire de celle-ci. La théorie générale de l'enquête par sondage assure que l'on peut utiliser les données provenant d'un échantillon aléatoire pour établir des estimations non biaisées des paramètres de la population finie et de leurs erreurs d'échantillonnage, indépendamment de la complexité de conception de l'échantillon. Cela suppose que l'échantillon soit un échantillon probabiliste, ce qui veut dire que chaque unité de la population a une probabilité positive connue d'être sélectionnée. Ces méthodes classiques d'estimation des paramètres de la population finie sont connues sous le nom de méthodes fondées sur la conception (ou sur la randomisation), car toute déduction se fait sur la base des propriétés de l'échantillon, à travers la distribution probabiliste. Il convient toutefois de noter que l'on ne peut généralement évaluer l'efficacité de différentes stratégies d'estimation (la conception d'un échantillon combinée à une formule d'estimation) que lorsque l'on dispose d'amples informations sur la population, ce qui n'est généralement pas le cas dans la pratique. Autrement dit, même les écrits classiques sur l'échantillonnage (par exemple, Cochran, 1977) s'appuient souvent sur les modèles pour justifier une méthode donnée d'échantillonnage ou d'estimation. Si les chiffres de la population suivent un simple modèle de régression, par exemple sous certaines hypothèses, l'estimateur des ratios est plus efficace qu'un simple estimateur d'expansion. Les méthodes fondées sur la conception sont donc souvent assistées par un modèle, mais pas fondées sur un modèle (ou dépendantes d'un modèle) : avec les méthodes de conception d'échantillon et d'estimation de statistiques descriptives assistées par un modèle, on n'a pas besoin de suppositions pour une estimation presque exempte de distorsion des paramètres d'une population finie.

10. L'approche de la conception d'un échantillon et de l'estimation fondée sur un modèle (ou théorie de prédiction) suppose que les chiffres de la population finie sont en fait des réalisations d'une distribution d'une super-population fondées sur un modèle hypothétique avec paramètres de super-population (modèle). Pour plus de renseignements et pour une analyse plus approfondie, voir Brewer et Mellor (1973), Hansen, Madow et Tepping (1983), Särndal, Swensson et Wretman (1992) et Valliant, Dorfman et Royall (2000). Par contraste avec les méthodes d'estimation assistée par un modèle, l'efficacité accrue obtenue par les méthodes d'estimation fondées sur un modèle est fonction de la validité du modèle utilisé. Autrement dit, si l'on a le moindre doute quant à la validité des hypothèses du modèle, la réduction apparente en carré moyen ne justifie pas nécessairement l'utilisation d'une analyse purement fondée sur un modèle. Hansen, Madow et Tepping (1983) en donnent un bon exemple, où le fait de ne supposer aucune ordonnée à l'origine dans un modèle de régression, même si cette ordonnée est en fait très proche de zéro, se traduit par une déduction non valable fondée sur un modèle.

11. Les enquêtes sont utilisées de plus en plus à des fins analytiques ainsi que descriptives. Souvent, leur conception tient compte de l'intention de les utiliser à des fins analytiques. Cela est dû au fait que les décideurs et les chercheurs s'intéressent,

dans les processus concernant les données brutes, à établir des modèles des relations entre les variables étudiées. Bien entendu, ces analyses supposent que l'on formule des hypothèses quant aux modèles. Le but d'une telle analyse est de confirmer la validité d'un modèle supposé et d'estimer les paramètres du modèle plutôt que ceux de la population finie. Ainsi, l'analyse est par essence fondée sur un modèle. Les déductions quant aux paramètres du modèle doivent, pratiquement par définition, être fondées sur les modèles étudiés.

12. Il convient toutefois de souligner que lorsque la population est très nombreuse et que les modèles choisis sont valables, il y a en réalité très peu de différence entre les paramètres du modèle et leurs homologues dans la population finie. Par exemple, si le modèle de régression linéaire type

$$Y_i = \alpha + \beta X_i + \varepsilon_i, \text{ avec } \varepsilon_i | X_i \underset{\text{ind}}{\sim} N(0, \sigma_i^2)$$

est valable et la population nombreuse, la valeur du coefficient type de régression de la population B (voir annexe) est très proche de la valeur du paramètre du modèle β , en raison du théorème de la limite centrale. Autrement dit, s'il est vrai que nous nous concentrerons ci-après sur l'estimation des paramètres du modèle, ces paramètres seront parfois remplacés par leurs homologues de la population finie. Pour simplifier la présentation, la plupart des exemples ci-après seront formulés sur la base de distributions d'univariates (c'est-à-dire d'une variable dépendante unique et d'une variable explicative unique). L'extrapolation des résultats aux cas de multivariates ne pose généralement pas de problème.

13. En résumé, les modèles hypothétiques font partie intégrante de l'analyse statistique. Pour analyser les données d'enquêtes par sondage, le choix d'un bon modèle convenant aux données est déterminant pour l'analyse. Il faut que les chercheurs et les analystes aient une bonne compréhension des modèles sur lesquels reposent les processus qu'ils désirent étudier avant d'appliquer les méthodes analytiques. Comme nous le verrons dans les sections suivantes, l'application de données d'échantillons de conception complexe exige une bonne compréhension du modèle sous-jacent et de la façon dont la complexité de conception peut influencer sur l'analyse

4. Effets de conception et leur rôle dans l'analyse des données d'un échantillon complexe

14. Les questions concernant les effets de conception et leur estimation ont été longuement traités aux chapitres VI et VII, principalement du point de vue de leur rôle dans la conception et l'estimation dans les enquêtes d'énumération. Dans ce chapitre, nous verrons qu'elles jouent aussi un rôle important dans l'analyse des données d'enquêtes par sondage complexes. L'idée est que, si l'on suppose que les hypothèses du modèle sont valables, on peut facilement obtenir des estimations non biaisées des paramètres du modèle et des estimations de la variance de ces estimations au moyen d'un simple échantillonnage aléatoire. Ces estimations et ces estimations de la variance servent de base à l'évaluation des hypothèses concernant les paramètres du modèle. Par exemple, avec un échantillonnage aléatoire simple et en prenant un modèle de régression simple

$$Y_i = \alpha + \beta X_i + \varepsilon_i, \text{ avec } \varepsilon_i | X_i \underset{\text{ind}}{\sim} N(0, \sigma_i^2)$$

l'estimateur de l'échantillon des moindres carrés ordinaires, b , est un estimateur non biaisé de β , et l'on dispose d'un estimateur non biaisé de sa variance, $v(b)$ [voir annexe]. Le test standard de l'hypothèse où $\beta = 0$ est donc fondé sur la statistique test $b/\sqrt{v(b)}$, en vertu du théorème de la limite centrale. Lorsque l'échantillon est de conception complexe, par exemple lorsqu'il s'agit d'un échantillon en grappe strati-

fié, l'estimateur b reste non biaisé par le modèle, si le modèle de régression reste valable et si la conception de l'échantillon n'est pas fonction des valeurs de Y_i (par exemple à l'inverse de la situation décrite aux figures XX.4-XX.6). Par cela, nous voulons dire qu'avec le modèle de régression spécifié, la valeur attendue de b est β , ce qui correspond à ce que l'on attend de la distribution des valeurs de Y_i au sein de la super-population. Comme nous le verrons à la section C.1, cela n'est plus nécessairement vrai si le modèle n'est plus valable. Toutefois, même si les hypothèses sur lesquelles repose le modèle restent valables, $v(b)$ n'est plus un estimateur valable de la variance de b du modèle et doit être modifié. Souvent, on peut calculer un estimateur direct de la variance correcte du modèle, en utilisant l'un des logiciels décrits au chapitre XXI, et l'utiliser à la place de $v(b)$. En l'absence d'un estimateur direct, on peut souvent obtenir un bon estimateur de l'effet lié à la conception, exprimé par le terme $d^2(b)$. Cet estimateur peut être utilisé pour modifier la statistique test $v(b)$ par $d^2(b) \times v(b)$. Nous présenterons ci-dessous d'autres utilisations spécifiques des effets de conception pour modifier les statistiques de test type pour d'autres applications

B. Approches fondamentales de l'analyse des données d'échantillons complexes

1. Utilisation des spécifications du modèle comme base de l'analyse

15. Une définition précise de modèle est indispensable à toute analyse. Les conséquences d'une mauvaise définition — aussi bien l'exclusion de variables explicatives pertinentes (ou l'inclusion de variables superflues) que l'utilisation d'une forme fonctionnelle inappropriée (par exemple, linéaire au lieu de quadratique) — sont connues et traitées dans des ouvrages courants. Elles peuvent se présenter sous la forme de distorsions dans l'estimation des paramètres du modèle (principalement de l'exclusion de variables pertinentes), de pertes d'efficacité (liées surtout à l'inclusion erronée de variables explicatives) et de tailles et de puissance modifiées dans les tests des hypothèses. Ces effets peuvent être aggravés lorsque la mauvaise définition a trait directement aux variables de conception de l'échantillon ou à des variables corrélées avec celles-ci. Quoi qu'il en soit, il ne faut pas perdre de vue que les variables de conception de l'enquête ne conviennent pas pour les objectifs de la recherche. De plus, leur inclusion dans le modèle analytique peut très bien ne pas être justifiée.

16. L'incorporation des variables de conception de l'enquête dans le modèle peut se faire selon deux approches essentielles. L'approche globale considère le modèle recherché comme se situant au niveau de la population et comme étant théoriquement indépendant de l'échantillon conçu pour obtenir les données. Selon cette approche, les variables de conception ne seraient incluses dans le modèle que si elles sont pertinentes en regard de l'analyse du sujet. Par exemple, disons que nous voulons expliquer la variable binaire employé/sans emploi par la variable explicative des années d'études, indépendamment de la localisation géographique. L'échantillon est stratifié, disons, par régions géographiques auxquelles différents modèles peuvent être applicables. Tel serait le cas même si l'on avait utilisé un échantillonnage aléatoire simple. De ce fait, la stratification peut être incluse dans le modèle (voir l'approche désagrégée examinée ci-après) pour refléter les variations régionales dans les relations entre les variables du modèle. Si, en revanche, la stratification et l'allocation de l'échantillon aux strates se font pour des raisons purement opérationnelles (convenance et coût), les pondérations de l'échantillon risquent de ne pas convenir au modèle de population. L'incorporation des pondérations d'échantillonnage dans une analyse autrement exempte d'effets liés aux strates entraîne une perte d'efficacité. Cependant, elle permet une interprétation

aisée d'un modèle exempt d'effets liés aux strates mais résistant à l'échec si certains de ces effets se produisent réellement.

17. L'approche désagrégée élargit le modèle de l'analyste non seulement aux variables étudiées mais aussi aux variables utilisées lors de la conception de l'enquête et à celles qui ont trait à la structure de la population considérée lors de la conception. Les variables de conception se rapportant à la stratification et à la mise en grappes sont incluses dans le modèle de manière à refléter la complexité de la structure de la population. En effet, dans l'exemple précédent, le modèle comporterait un ensemble différent de coefficients (une ordonnée à l'origine et une pente) pour chaque strate géographique. Avec l'approche désagrégée, on tient pleinement compte de la conception de l'échantillon, et l'on suppose que toutes les variables de conception sont à juste titre incluses dans le modèle. Le grand nombre de paramètres qu'il faut estimer selon cette approche peut être source de difficultés et produire des estimations moins précises que celles de modèles agrégés plus parcimonieux. L'approche désagrégée ne convient que lorsque l'analyste estime que le modèle imaginé convient à son propos.

18. Le choix de l'approche appropriée, agrégée ou désagrégée, dépend de ce que vise l'analyste. L'approche agrégée convient mieux à l'étude des facteurs qui touchent la population tout entière et, à ce titre, peut être plus utile pour évaluer les décisions de politique nationale. L'approche désagrégée prête davantage à l'étude des micro-effets et des effets des décisions d'intérêt local ou sectoriel. Pour d'autres exemples et une analyse plus approfondie, voir Skinner, Holt et Smith (1989) et Chambers et Skinner (2003).

2. Relations possibles entre le modèle et la conception de l'échantillon : conceptions informatives et non informatives

19. Il importe de faire une distinction entre les conceptions d'échantillon informatives et non informatives lorsque l'on analyse les données d'une enquête complexe. Une fois qu'un modèle a été imaginé, l'analyste doit se demander si, après conditionnement sur les covariates du modèle, les probabilités de sélection de l'échantillon sont liées aux valeurs de la variable de réponse. Le processus d'échantillonnage est informatif si la distribution conditionnelle commune des observations du modèle pour l'échantillon, compte tenu des valeurs des covariates du modèle, diffère de leur distribution conditionnelle au sein de la population. La conception de l'échantillon n'est non informative (ou négligeable) que lorsque ces distributions sont identiques, auquel cas on peut employer des méthodes analytiques courantes comme si les observations provenaient d'un échantillon aléatoire simple. Lorsque la conception de l'échantillon est informative, le modèle valable pour les données de l'échantillon diffère du modèle de population. Si l'on ne tient pas compte du processus d'échantillonnage en pareil cas, on risque d'obtenir des estimateurs biaisés et de fausser l'analyse de la même manière que lorsque l'on exclut des variables du modèle dans une analyse conventionnelle. Il convient de noter que l'inclusion à juste titre de variables liées à la conception dans le modèle assure une conception non informative.

20. Le fait d'inclure toutes les variables liées à la conception dans le modèle pose deux problèmes importants. Premièrement, on risque de ne pas savoir exactement quelles variables ont été utilisées dans la conception ou, si on le sait, on peut ne pas connaître leurs valeurs. Même lorsque les variables de conception sont identifiées et mesurées, il se peut que l'analyste ne connaisse pas la forme exacte de la relation (par exemple, linéaire ou exponentielle) entre elles et la variable étudiée. En effet, si la conception est stratifiée, il faut vérifier si une relation de régression a des pentes et des ordonnées à l'origine différentes pour les différentes strates.

21. Deuxièmement, lorsque les variables de conception sont correctement incluses dans le modèle, les estimations qui en résultent risquent d'être sans grand intérêt pour l'analyste car les variables ajoutées ne présentent pas vraiment d'intérêt pour le sujet traité (voir la discussion sur les analyses agrégée et désagrégée). Cela implique que l'on ne peut pas toujours traiter l'effet de la conception d'un échantillon complexe sur l'analyse uniquement en modifiant le modèle. Dans la section suivante, nous examinerons à la fois comment modifier les méthodes analytiques courantes de manière à tenir compte d'une conception complexe et comment définir des méthodes d'estimation et d'analyse propres à la conception retenue.

3. Problèmes liés à l'utilisation de logiciels courants pour l'analyse d'échantillons complexes

22. L'utilisation quasi universelle de logiciels courants pour l'analyse statistique a engendré un abus généralisé des saines pratiques statistiques, qui est souvent encore aggravé lorsque l'on analyse les données d'une enquête par sondage complexe.

23. Malheureusement, les avantages que présentent les logiciels statistiques pour ce qui est de faciliter l'analyse s'accompagnent du risque d'effectuer une analyse sans comprendre les principes sur lesquelles elle repose. Cela pose un grave problème pour les travaux quantitatifs, en particulier dans les sciences sociales. Le problème est encore accentué par le fait que les logiciels les plus courants traitent les données comme si elles provenaient d'un simple échantillonnage aléatoire. Comme on l'a souligné précédemment, cela peut entraîner de graves distorsions lorsque la conception est informative. Néanmoins, si l'on prend les précautions nécessaires, on peut souvent adapter le logiciel normalisé de manière à saisir ou à présenter approximativement l'effet d'une conception complexe. En particulier, la procédure *SURVEYREG* des versions les plus récentes de SAS® (versions 8 et 9) présente une analyse de régression qui tient compte de la conception de l'échantillon d'une manière semblable à celle décrite ci-dessous [voir An et Watts (2001)].

24. Par exemple, considérons le modèle de régression (hétéroscédastique) linéaire défini par :

$$Y_i = \alpha + \beta X_i + \varepsilon_i, \text{ avec } \varepsilon_i | X_i \underset{\text{ind}}{\sim} N(0, \sigma_i^2).$$

Généralement, les programmes informatiques courants calculent b , l'estimateur des moindres carrés ordinaires [*ordinary least squares* (OLS)] de β , ou l'estimateur des moindres carrés généralisés [*generalized least squares* (GLS)], b_G , où les sommes et les produits sont pondérés par les inverses de σ_i^2 , dont les valeurs (ou valeurs relatives) sont censées être connues (voir annexe). Ces deux estimateurs sont des estimateurs non biaisés du paramètre β , si le modèle est valable, mais b_G est un estimateur plus efficace dans le cas hétéroscédastique. Les programmes courants offrent aussi des estimateurs des variances de l'estimateur OLS, $v(b)$, et de l'estimateur GLS, $v(b_G)$, qui sont chacun non biaisés avec le modèle approprié [le modèle hétéroscédastique dans le cas de $v(b)$].

25. Souvent, on peut avoir des doutes quant à la validité du modèle; aussi, au lieu d'estimer β , il peut être préférable d'estimer l'homologue de β , que nous appellerons B (voir annexe). Bien que b (l'estimateur OLS) soit non biaisé pour $\beta \approx B$, en général, il n'est pas exempt de distorsions liées à la conception. Dans des conditions appropriées, l'estimateur pondéré en fonction de l'échantillon, b_W (Horvitz-Thompson), où les produits et les carrés sont pondérés par les inverses des probabilités d'inclusion, est à la fois conforme à la conception et non biaisé par rapport au modèle. En outre, on peut obtenir b_W à partir des options de régression pondérées de nombreux programmes en

utilisant les valeurs de w_i comme pondérations. On peut aussi l'obtenir par régression non pondérée des variables $Y_i/\sqrt{\pi_i}$, $X_i/\sqrt{\pi_i}$, en remplaçant l'ordonnée à l'origine par $1/\sqrt{\pi_i}$. Il importe toutefois de souligner que, dans un cas comme dans l'autre, les estimations de la matrice variance-covariance donnée par la plupart des programmes courants est inexacte, tant comme estimateur de l'erreur carrée moyenne de conception que comme estimateur de la variance du modèle, sauf dans des cas exceptionnels.

26. En résumé, l'utilisation de programmes logiciels courants qui ne tiennent pas compte de la complexité de conception des enquêtes est à éviter, à moins que l'on ne puisse déterminer que cette complexité est sans effet sérieux sur l'estimation. Cela est possible par une bonne utilisation de ces logiciels. (Voir l'exemple de la section C.2.) L'utilisation de logiciels traitant spécifiquement d'échantillons de conception complexe est recommandée (voir chapitre XXI).

C. Analyse de régression et modèles linéaires

1. Effet des variables de conception ne figurant pas dans le modèle et des estimateurs de régression pondérée

27. L'analyse de régression et la modélisation linéaire sont des applications très courantes où des modèles types conçus pour des échantillons aléatoires simples sont régulièrement appliqués aux données provenant d'enquêtes par sondage complexes. Comme on l'a déjà souligné, cela peut souvent conduire à des analyses et des conclusions erronées. L'un des principaux moyens de prévenir les erreurs est d'identifier les variables qui déterminent la conception de l'échantillon ou qui l'influencent, afin de pouvoir les inclure dans le modèle. Toutefois, comme nous l'avons vu, même lorsque ces variables sont identifiées, leur inclusion dans le modèle n'est pas nécessairement justifiée du point de vue du sujet traité. Dans la présente sous-section, nous étudierons les conséquences du fait de ne pas inclure ces variables dans le modèle et nous examinerons les possibilités de modifier ces estimateurs de manière à tenir compte de la complexité de conception. Pour faciliter la présentation, nous examinerons le cas d'une variable dépendante unique, Y (désignée pour des raisons techniques par le terme X_1 dans cette section), où le modèle étudié a une seule variable explicative (X_2), et où il n'y a qu'une variable de conception (X_3). Le modèle étudié est donc $E(X_1) = \mu_1 + \beta_{12}(X_2 - \mu_2)$, où β_{12} est le paramètre étudié, plutôt que le modèle complet, qui comprend la variable de conception, X_3 . Pour la formule utilisée dans cette sous-section, voir l'annexe.

28. Dans les conditions assez générales spécifiées dans Nathan et Holt (1980), l'estimateur standard OLS pour β_{12} , $b_{12} = s_{12}/s_2^2$, peut être à la fois biaisé par le modèle, selon X_3 et selon l'échantillon, S , et biaisé de façon inconditionnelle. Les expressions correspondant à l'attente du modèle conditionnel et son attente inconditionnelle (modèle et conception conjoints) montrent qu'en général b_{12} est asymptotiquement biaisé, à moins que ρ_{23} , la corrélation entre X_2 et X_3 , soit un estimateur non biaisé de sa vraie variance. On peut montrer que cette deuxième condition tient asymptotiquement pour un grand nombre d'échantillons de conception d'égale probabilité (*epsem*), mais rarement pour des échantillons de probabilité inégale (par exemple, des échantillons de conception stratifiée non proportionnelle).

29. On peut utiliser un estimateur corrigé, asymptotiquement non biaisé, fondé sur l'estimateur de probabilité maximum dans des conditions normales, $\hat{\beta}_{12}$, à la place de l'estimateur OLS. Les expressions des variances de b_{12} et de $\hat{\beta}_{12}$ sont données par Nathan et Holt (1980). Il faut noter que l'estimateur habituel de la variance de b_{12} , $v(b_{12})$, risque d'être loin d'être non biaisé, même lorsque b_{12} est un estimateur régulier de β_{12} . Il peut se produire que les valeurs de ε_i ne soient pas indépendantes et identiquement distribuées parmi les observations de l'échantillon.

30. Ni l'estimateur b_{12} ni $\hat{\beta}_{12}$ ne dépend de la conception de l'échantillon, mais tel n'est pas le cas de leurs propriétés. Les informations sur la conception de l'échantillon peuvent contribuer utilement à améliorer ces estimateurs, soit lorsque l'on ne dispose pas d'informations sur les valeurs de la variable de conception X_3 pour l'ensemble de la population (auquel cas, on peut utiliser S_3^2 pour l'estimation), soit lorsque l'analyste veut assurer la solidité des écarts par rapport au modèle. Pour ce faire, on peut utiliser des estimateurs pondérés pour l'échantillon fondés sur l'estimation de Horvitz-Thompson pour chacune des composantes des estimateurs non pondérés. On peut remplacer les cas d'échantillons non pondérés par leurs versions pondérées dans les expressions pour b_{12} et pour $\hat{\beta}_{12}$ afin d'obtenir les estimateurs pondérés b_{12}^* et $\hat{\beta}_{12}^*$.

31. On notera que l'on peut utiliser b_{12}^* quand la variance de population de X_3 , S_3^2 , est inconnue, mais que l'on ne peut pas utiliser $\hat{\beta}_{12}$ dans cette situation. On voit facilement que, dans des conditions relativement générales, ces deux estimateurs sont compatibles avec le paramètre de la population finie, $\hat{\beta}_{12}^*$.

32. Nathan et Holt (1980) présentent des comparaisons empiriques de la performance de ces quatre estimateurs pour une population de $N = 3\,850$ exploitations, pour lesquelles on dispose des données sur les terres de culture (X_1), la superficie totale (X_2) et la valeur totale de la production de l'année précédente (X_3). Les exploitations ont été stratifiées sur la base des valeurs de X_3 , ce qui a donné six strates comprenant respectivement 563, 584, 854, 998, 696 et 155 exploitations. On a utilisé les cinq conceptions d'échantillon suivantes pour choisir des échantillons de taille $n = 400$ (voir tableau XX.1):

- (A) Échantillonnage aléatoire simple;
- (B) Échantillonnage aléatoire simple proportionnel stratifié;
- (C) Échantillonnage aléatoire simple stratifié à taille fixe;
- (D) Échantillonnage aléatoire simple stratifié avec allocation plus élevée que proportionnelle aux strates ayant des valeurs de X_3 élevées (25, 30, 60, 80, 130, 75);
- (E) Échantillonnage aléatoire simple stratifié avec allocation en U allocation (100, 80, 20, 20, 80, 100).

Tableau XX.1

Distorsion et carré moyen de l'estimateur des moindres carrés ordinaires et variances des estimateurs non biaisés pour une population de 3 850 exploitations, déterminés à partir de diverses conceptions d'enquêtes

Conception d'enquête	$E(b_{12}) - \beta_{12}$	MSE (b_{12})	$V(\hat{\beta}_{12}^*)$	$V(b_{12}^*)$	$V(\hat{\beta}_{12}^*)$
A	0-000	0-000214	0-000197	0-000226	0-000197
B	0-000	0-000200	0-000198	0-000222	0-000196
C	0-031	0-001102	0-000160	0-000222	0-000196
D	0-027	0-000879	0-000163	0-000220	0-000195
E	0-042	0-001877	0-000152	0-000225	0-000196

Source : Nathan et Holt (1980); tableau 1.

33. Les résultats indiquent la distorsion de b_{12} pour les conceptions non *epsem* (C, D, E), tandis que les autres estimateurs sont soit persistants selon la conception ou selon le modèle (ou selon les deux). Ils montrent aussi l'avantage que présente $\hat{\beta}_{12}^*$ par rapport aux estimateurs pondérés pour toutes les conceptions considérées. Cela tient,

bien que les hypothèses utilisées pour l'ensemble du modèle paraissent ne pas tenir pour la population. Toutefois, lorsque S_3^2 est inconnu, le facteur b_{12}^* , moins efficace mais toujours persistant, reste un estimateur raisonnable.

34. En résumé, lorsque les données sont fondées sur des conceptions de probabilité inégale, il est judicieux de considérer à la fois les estimateurs du maximum de vraisemblance pondérés et non pondérés, plutôt que les simples estimateurs OLS. L'estimateur non pondéré semble être plus efficace. Toutefois, dans beaucoup d'applications, l'analyste ne dispose pas des informations nécessaires pour calculer les estimateurs du maximum de vraisemblance, et des estimateurs pondérés selon l'échantillon, moins efficaces mais persistants, sont en fait régulièrement utilisés [voir Korn et Graubard (1999)].

2. Évaluation de l'effet de la conception sur l'analyse de régression

35. Nombreux sont les analystes qui préfèrent utiliser les coefficients de simples estimateurs de régression pondérés ou non pondérés, que l'on peut tirer de logiciels standard, plutôt que les estimateurs modifiés proposés à la section C.1. Nous avons vu que le simple estimateur OLS est persistant lorsque la conception est non informative ou lorsque l'effet de la conception est négligeable, mais que son homologue pondéré est préférable dans les cas contraires. DuMouchel et Duncan (1983) ont proposé un test simple, fondé sur les ensembles logiciels courants, pour décider si les pondérations devraient être utilisées lorsqu'on se fonde sur les données d'un échantillon non en grappe. Considérons le cas univariate à une seule variable explicative. L'extension au cas multivariate est automatique. Si $\hat{\Delta} = b_w - b$, il s'agit de tester l'hypothèse : $\Delta = E(\hat{\Delta}) = 0$. DuMouchel et Duncan ont montré que le test pour $\Delta = 0$ est le même que le test pour $\gamma = 0$, avec le modèle $Y_i = \alpha + \beta X_i + \gamma Z_i + \varepsilon_i$, où $Z_i = w_i X_i$ et $\varepsilon_i | X_i \sim N(0, \sigma_i^2)$. Les auteurs ont donné un exemple numérique pour le cas multivariate ^{ind} portant sur un sous-ensemble de données de l'Étude de la dynamique des revenus d'un groupe du Centre de recherche sur les enquêtes de l'Université du Michigan. L'échantillon de 658 personnes a été sélectionné selon diverses probabilités, ce qui s'est traduit par des pondérations comprises entre 1 et 83. Le modèle final utilisé pour expliquer le niveau d'instruction comprenait une constante et 17 variables explicatives, telles que le niveau d'instruction des parents, le revenu, l'âge, la race, l'emploi et les interactions. Le tableau suivant de l'analyse de la variance (ANOVA) a été obtenu :

Tableau XX.2

Tableau ANOVA comparant les régressions pondérées et non pondérées

Source	df	Somme des carrés	Carré moyen	F	Signification
Régression	17	730,6	43,0	17,35	<0,0001
Pondérations	18	43,3	2,5	0,97	0,494
Erreur	622	1 542,2	2,5		
Total	657	2 315,9			

36. Prises ensemble, les 18 variables correspondant à Z_i (les 17 variables explicatives et la constante, chacune multipliée par w_i) ont une valeur F de 0,97 et un niveau de signification de 0,494 seulement. Autrement dit, une régression non pondérée est justifiée, même si elle risque d'entraîner une perte de puissance.

37. En général, les analystes peuvent montrer la même appréhension à accepter l'hypothèse nulle, $\Delta = E(\hat{\Delta}) = 0$, quand elle est fautive qu'à la rejeter quand elle est

vraie. De ce fait, ils peuvent décider d'effectuer une analyse pondérée (avec le logiciel approprié) ou de développer un modèle moins parcimonieux lorsque le niveau de signification est considérablement plus élevé que la norme 0,05. Dans l'exemple ci-dessus, le niveau de signification est très proche de 0,05, ce qui semble indiquer que l'on peut ignorer les pondérations. Cependant, dans une version précédente du modèle, le niveau de signification de Z_i était de 0,056, et DuMouchel et Duncan ont ajouté quelques termes d'interaction. Les résultats finals sont ceux qui figurent dans le tableau.

38. Le test de DuMouchel-Duncan décrit ci-dessus suppose que les termes ε_i sont indépendants et répartis de façon identique. Souvent, les données d'enquêtes proviennent d'échantillons à phases multiples. Lorsque les valeurs ε_i des observations ont une hétéroscédasticité inconnue quelle que soit la conception, ce test ne convient pas. Néanmoins, l'analyste peut estimer que le fait d'utiliser les pondérations de l'échantillon ajoute une variance inutile aux estimations produites. On peut utiliser un test effectué par Wald selon les règles proposées par Fuller (1984). Dans la pratique, cela implique que l'on utilise un logiciel tel que SAS/SURVEYREG et que l'on entre deux fois chaque donnée, une fois avec la pondération de l'échantillon fixée à 1 et une fois cette pondération fixée à sa valeur réelle.

39. Pfeffermann et Sverchkov (1999) ont proposé l'utilisation d'un autre ensemble de pondérations lorsque le modèle linéaire est correct et que les erreurs sont indépendantes et identiquement distribuées mais que la conception de l'échantillon est informative. Le test décrit ci-dessus peut servir à évaluer ces pondérations par rapport à celles de l'échantillon. Pour un examen plus approfondi du rôle des pondérations de l'échantillon dans la modélisation des données d'enquêtes, voir Pfeffermann (1993) et Korn et Graubard (1999).

3. Modèles à niveaux multiples et échantillon de conception informative

40. On assiste depuis peu à une utilisation accrue des modèles à niveaux multiples pour l'analyse de données de populations à structure hiérarchique complexe. Par exemple, dans la plupart des enquêtes sur les ménages, les individus intégrés aux familles sont des unités étudiées, et l'on s'intéresse à la fois aux relations entre individus et entre ménages. Il existe des structures hiérarchiques semblables pour les enquêtes sur les élèves des écoles et les employés d'entreprise.

41. On peut facilement étendre les modèles linéaires habituels à niveau unique de manière à tenir compte d'une hiérarchie en utilisant des modèles mixtes (à effets aléatoires et fixes) avec une structure d'erreur reflétant la configuration hiérarchique. Par exemple, le modèle connu sous le nom de modèle aléatoire avec ordonnée à l'origine peut être donné par la formule suivante (pour une variable explicative unique) :

$$y_{ij} = \beta_{oi} + \beta x_{ij} + \varepsilon_{ij}; \quad \varepsilon_{ij} | x_{ij} \sim N(0, \sigma_\varepsilon^2); \quad (i = 1, \dots, N; j = 1, \dots, M_i)$$

où y_{ij} est la variable de sortie pour l'unité de premier niveau j (disons un individu) au sein d'une unité de deuxième niveau (disons un ménage), x_{ij} est une variable explicative connue et β un paramètre inconnu. L'ordonnée à l'origine, β_{oi} , est ici une variable aléatoire, qui est à nouveau modélisée sous la forme :

$$\beta_{oi} = \alpha + \gamma z_i + u_i; \quad u_i | z_i \sim N(0, \sigma_u^2); \quad (i = 1, \dots, N)$$

où z_i est une variable explicative de l'unité de deuxième niveau et α et γ sont des paramètres inconnus.

42. Avec un simple échantillonnage aléatoire, on peut analyser les modèles de ce type en utilisant des extensions directes de la théorie du modèle linéaire à niveau unique. Malheureusement, on ne dispose pas de formes fermées pour les estimations des paramètres du modèle (α , β , γ , σ_ε^2 , σ_u^2 pour le modèle ci-dessus). On utilise donc une procédure itérative, les moindres carrés généralisés itérés [*Iterated Generalized Least Squares* (IGLS)]. Cette procédure donne des estimations qui convergent vers des solutions au maximum de vraisemblance. De la sorte, on ne peut employer dans ce cas les méthodes fermées d'adaptation des moindres carrés pondérés pour tenir compte de la conception de l'échantillon. Une version d'IGLS pondérée en fonction de l'échantillon (PWIGLS), qui pondère les équations d'estimations des premier et deuxième niveaux à l'aide de pondérations appropriées fondées sur les probabilités de sélection, a été mise au point pour obtenir des estimateurs persistants des paramètres [pour plus de détails, voir Pfeffermann *et al.* (1998)].

43. Plus récemment, Pfeffermann, Moura et Silva (2001) ont proposé une approche dépendante du modèle (purement fondée sur le modèle) pour une analyse à niveaux multiples qui tient compte de l'échantillonnage informatif. Cette approche avait pour but d'extraire le modèle hiérarchique valable pour les données de l'échantillon comme fonction du modèle de population et des probabilités d'inclusion d'un échantillon de premier ordre et d'adapter le modèle d'échantillon en utilisant les techniques d'estimation modernes. Les probabilités de sélection deviennent de nouvelles variables de sortie à modéliser de manière à renforcer la performance des estimateurs. On trouvera de plus amples renseignements qui débordent le cadre du présent chapitre dans Pfeffermann, Moura and Silva (2001). Une expérience de simulation qui suit de près la conception de l'étude d'évaluation de l'éducation de base à Rio de Janeiro de 1996 indique que les résultats de l'application de la méthode proposée sont prometteurs.

D. Analyse des données catégoriques

1. Modifications des tests de chi-carré en tests mesurant la bonne adaptation et l'indépendance

44. Les premières tentatives d'évaluation des effets de la complexité de conception d'un échantillon sur l'analyse de données catégoriques (données telles que chaque point tombe dans l'une d'un certain nombre fini de catégories ou de cellules) avaient trait à des modifications des tests de chi-carré qui sont couramment utilisés soit pour évaluer la bonne adaptation entre la distribution d'une variable catégorique unique et une distribution hypothétique soit pour déterminer l'indépendance entre deux variables catégoriques. Bien que plusieurs versions modifiées des tests de chi-carré aient été proposées par divers auteurs pour les données d'un échantillonnage aléatoire simple stratifié proportionné, l'effet de ces modifications est généralement très faible dans la pratique. Ainsi, dans une étude de statistiques découlant de tests de chi-carré modifiés effectuée en Israël sur huit ensembles de données provenant d'échantillons stratifiés proportionnés et présentées au tableau XX.3 [de Kish et Frankel (1974)], aucune des statistiques d'itération finales ne s'écartait de plus de 4 % de celles qui auraient été obtenues selon des hypothèses d'échantillonnage aléatoire simple [*simple random sampling* (SRS)], et la plupart d'entre elles s'en écartaient de moins de 1 %.

45. Bien que l'impact de la conception sur l'analyse de données catégoriques émanant d'un échantillonnage aléatoire simple stratifié proportionné soit généralement faible, il en va souvent différemment de l'échantillonnage en grappe, comme le montre l'étude séminale de Rao et Scott (1981). Pour les tests de bonne adaptation, ils ont montré que, dans l'hypothèse nulle, la statistique habituelle de chi-carré, X^2 , est

distribuée de façon asymptotique comme une somme pondérée de variables aléatoires $k - 1$ indépendantes χ^2_1 (c'est-à-dire de carrés normaux). Les pondérations sont les valeurs *eigen* d'une matrice D (voir annexe). La matrice peut être perçue comme une extension multivariante naturelle de l'effet de conception pour les statistiques univariates (voir chapitres VI et VII). Ses valeurs *eigen*, λ^2_{0i} , sont appelées effets de conception généralisés, et l'on peut montrer qu'elles sont les effets de conception pour certaines combinaisons linéaires des effets de conception d_i^2 , de p_i (= la proportion estimée de la population de catégorie i). On peut obtenir une statistique chi-carré modifiée, X_c^2 , en divisant la statistique standard X^2 par la moyenne des estimations de ces effets de conception généralisés, appelés $\hat{\lambda}^2$. Cette modification n'exige qu'une connaissance des effets de conception des estimations de cellules. Bien que X_c^2 n'ait pas de distribution asymptotique χ^2_{k-1} dans l'hypothèse nulle, elle a la même valeur asymptotique attendue que χ^2_{k-1} (c'est-à-dire, $k - 1$), mais avec une forte variance. Il se trouve que l'on peut utiliser X_c^2 empiriquement pour évaluer la bonne adaptation en comparant la valeur de cette statistique à la valeur critique de χ^2_{k-1} . C'est ce que montre le tableau XX.4 [de Rao et Scott (1981)], qui présente les tailles véritables des tests fondés respectivement sur X^2 et sur X_c^2 pour six éléments de données de l'Enquête générale de 1971 sur les ménages du Royaume-Uni de Grande-Bretagne et d'Irlande du Nord. Cette enquête avait une conception stratifiée à trois phases.

Tableau XX.3
Ratios de trois tests de chi-carré itérés à des tests de SRS^a

Ensemble de données	Numéros de strates	Rang \times colonnes	Taille de l'échantillon	Trois tests de Nathan					
				Première itération			Dernière itération		
				χ^2	X_1^2	G	χ^2	X_1^2	G
1	4	3 \times 3	845	1.028	0.992	1.017	1.004	1.004	1.005
2	4	3 \times 3	821	1.088	0.963	1.043	0.999	1.003	1.001
3	4	3 \times 3	491	1.740	0.707	1.406	1.011	1.001	1.009
4	4	3 \times 3	2 528	1.095	0.959	1.049	1.003	1.005	1.003
5	6	2 \times 4	500	1.079	0.967	1.040	1.004	1.003	1.003
6	3	2 \times 2	120	1.013	0.967	1.009	1.008	0.969	1.007
7	5	2 \times 2	269	1.076	0.989	1.043	1.011	1.015	1.011
8	2	2 \times 4	81	1.368	0.889	1.186	1.029	1.037	1.029

Source : Adapté des données de Nathan (1972).

^a Huit tableaux de contingences fondés sur des échantillons stratifiés proportionnés émanant d'Israël : n^{os} 1-4 : épargne; n^o 5 : attitudes; n^o 6 : données d'hôpitaux; n^o 7 : médicaments pour volailles; et n^o 8 : expériences de perception.

Tableau XX.4
Estimation des tailles asymptotiques de tests fondés sur et sur X_c^2 de certains éléments de l'Enquête générale de 1971 sur les ménages du Royaume-Uni de Grande-Bretagne et d'Irlande du Nord; la taille nominale est de 0,05

Variable	k	m	$\hat{\lambda}^2$	Taille (χ^2)	taille (X_c^2)
G1 : Âge du bâtiment	3	33,1	3,42	0,41	0,05
G2 : Type de propriété	3	33,4	2,54	0,37	0,06
G3 : Type de logement	4	27,7	2,17	0,30	0,06
G4 : Nombre de pièces	10	34,6	1,19	0,14	0,06
G5 : Revenu hebdomadaire brut du ménage	6	26,6	1,14	0,10	0,06
G6 : Âge du chef de famille	3	34,6	1,26	0,10	0,05

Ces résultats montrent que l'utilisation de la statistique standard du chi-carré, X^2 , peut être très trompeuse, tandis que la statistique modifiée, X_c^2 , donne de très bons résultats.

46. On obtient des résultats similaires quand on effectue un test d'indépendance sur un tableau de contingences à deux dimensions. Pour un tableau de contingences ayant r colonnes et c rangs, l'hypothèse nulle étudiée est $H_0 : h_{ij} = P_{ij} - P_i + P_j = 0$ ($i = 1, \dots, r; j = 1, \dots, c$), où p_{ij} est la proportion de la population de la $(i, j)^e$ cellule et P_{i+} , P_{+j} sont les totaux marginaux. La statistique habituelle de chi-carré pour les données d'un échantillon aléatoire simple, X_1^2 , est distribuée de façon asymptotique comme chi-carré avec $b = (r - 1)(c - 1)$ degrés de liberté dans l'hypothèse nulle. Cela n'est pas nécessairement vrai lorsque l'échantillon est de conception complexe. En fait, la distribution asymptotique de X_1^2 est une somme pondérée de b variables aléatoires indépendantes χ_1^2 comme dans le cas du test de bonne adaptation.

47. On peut construire une statistique de Wald généralisée en se fondant sur l'estimation de la matrice complète variance-covariance des estimations, $\hat{h}_{ij} = p_{ij} - p_{i+}p_{+j}$ [voir détails dans Rao et Scott (1981)]. Heureusement, une correction de premier ordre, qui ne nécessite que les estimations des variances de \hat{h}_{ij} , $\hat{v}(\hat{h}_{ij})$, semble être une approximation adéquate. La statistique modifiée est définie comme $X_{I(C)}^2 = X_1^2 / \hat{\delta}^2$, où $\hat{\delta}^2$ est une moyenne pondérée des effets de conception estimés de \hat{h}_{ij} . Lorsque l'on ne dispose pas d'estimations de ces effets, comme c'est souvent le cas avec l'analyse secondaire de données publiées, on peut obtenir une autre modification en remplaçant $\hat{\delta}^2$ par $\hat{\lambda}^2$, moyenne pondérée des effets de conception estimés des proportions des cellules, \hat{d}_{ij} . L'adéquation de ces approximations dépend pour beaucoup de la variance relative des effets de conception. On dispose d'une correction de second ordre lorsque la variance relative est élevée.

Tableau XX.5

Dimensions asymptotiques estimatives de tests fondés sur X_1^2 , $X_1^2/\hat{\delta}^2$, et sur $(X_1^2/\hat{\lambda}^2)$ pour une classification croisée de certaines variables tirées de l'Enquête générale de 1971 sur les ménages du Royaume-Uni de Grande-Bretagne et d'Irlande du Nord; la taille nominale est de 0,05

Classification croisée	$r + c$	$\hat{\delta}^2$	$\hat{\lambda}^2$	Taille ($X_1^2/\hat{\delta}^2$)	Taille ($X_1^2/\hat{\lambda}^2$)	Taille ($X_1^2/\hat{\lambda}^2$)
G1 × G2	2 × 2	1,99	3,18	0,16	0,05	0,01
G1 × G3	2 × 3	1,97	2,36	0,22	0,05	0,03
G1 × G4	2 × 3	1,24	1,98	0,09	0,05	0,01
G1 × G5	2 × 6	0,91	1,23	0,04	0,05	0,02
G1 × G6	2 × 3	0,97	1,75	0,05	0,05	0,01
G2 × G3	2 × 3	1,94	2,49	0,21	0,05	0,03
G2 × G4	2 × 3	1,41	1,86	0,12	0,05	0,02
G2 × G5	2 × 6	1,02	1,18	0,06	0,05	0,03
G2 × G6	2 × 3	1,13	1,61	0,08	0,05	0,02
G3 × G4	3 × 3	1,26	1,72	0,11	0,05	0,01
G3 × G5	3 × 6	0,93	1,14	0,03	0,05	0,02
G3 × G6	3 × 3	0,96	1,51	0,05	0,05	0,01
G4 × G5	3 × 6	0,94	1,05	0,05	0,05	0,03
G4 × G6	3 × 3	0,93	1,21	0,04	0,05	0,02
G5 × G6	6 × 3	0,85	0,94	0,03	0,05	0,04

48. Le tableau XX.5 présente les effets empiriques pour 15 tableaux de contingences à deux dimensions fondés sur les données de l'Enquête générale sur les ménages du Royaume-Uni de Grande-Bretagne et d'Irlande du Nord [de Rao et Scott (1981)]. Ces effets indiquent une fois encore que : a) la statistique chi-carré non corrigée, X^2 , donne souvent des résultats très médiocres; b) la statistique corrigée, $X^2_{I(C)}$, fondée sur $\hat{\delta}^2$, atteint presque exactement la taille nominale; et c) la statistique corrigée fondée sur $\hat{\lambda}^2$ dévie dans le sens de la prudence.

2. Généralisations pour les modèles log-linéaires

49. Rao et Scott (1984) ont généralisé les résultats ci-dessus pour tableaux à deux dimensions au modèle log-linéaire utilisé dans l'analyse de tableaux à dimensions multiples. Soit π le vecteur- T des proportions des cellules de population, π_i , dans le tableau à dimensions multiples, avec $\sum_1^T \pi_i = 1$ (par exemple, $T = 4$ pour un tableau 2×2). Soit M_1 le modèle log-linéaire saturé (qui comprend toutes les interactions possibles). Nous voulons tester l'hypothèse selon laquelle un sous-modèle réduit, M_2 , est suffisant. Soit $\hat{\pi}$ le pseudo-estimateur de vraisemblance maximum de π dans le modèle M_1 . Il est défini comme la solution de l'estimation de l'échantillon des équations de vraisemblance du recensement (celles qui auraient été obtenues sur la base des données de la population) et fondé sur un estimateur de π conforme à la conception dans la conception de l'enquête (voir annexe). De même, soit $\hat{\pi}$ le pseudo-estimateur de vraisemblance maximum de $\hat{\pi}$ dans M_2 . La statistique standard chi-carré de Pearson pour tester H_0 , sur la base de $\hat{\pi}$, et de $\hat{\pi}$, n'a pas généralement de distribution asymptotique de chi-carré dans l'hypothèse nulle. Ce cas est similaire à celui d'un tableau à deux dimensions, dans la mesure où la distribution asymptotique statistique standard de chi-carré de Pearson est une somme pondérée de u variables aléatoires indépendantes χ^2 avec des pondérations δ_i^2 , qui sont les valeurs *eigen* d'une matrice générale des effets de conception (pour plus de détails, voir annexe).

50. Pour tenir compte de la complexité de conception, on propose des statistiques chi-carré modifiées, $X^2/\hat{\delta}^2$, $X^2/\hat{\lambda}^2$ et X^2/\hat{d}^2 . Ici $\hat{\delta}^2$ est la moyenne des valeurs *eigen* estimatives, $\hat{\lambda}^2$ est la moyenne des effets de conception estimatifs de X^2/\hat{p} et \hat{d}^2 est la moyenne des effets de conception estimatifs des cellules (pour plus de détails, voir annexe). Il convient de noter que $\hat{\lambda}^2$ et \hat{d}^2 ne dépendent pas de l'hypothèse nulle, H_0 , ce qui n'est pas le cas de $\hat{\delta}^2$. Par ailleurs, \hat{d}^2 nécessite que l'on ne connaisse que les effets de conception des cellules, au même titre que $\hat{\lambda}^2$ lorsque M_1 est le modèle saturé.

51. Dans le cas important de modèles admettant des solutions implicites pour $\hat{\pi}$ et pour $\hat{\pi}$, Rao et Scott (1984) montrent que l'on peut calculer $\hat{\delta}^2$ en ne connaissant que les effets de conception des cellules et ceux de leurs termes marginaux. Par exemple, dans l'hypothèse d'une complète indépendance dans un tableau à trois dimensions $I \times J \times K$, $H_0 : \pi_{ijk} = \pi_{i++}\pi_{+j+}\pi_{++k}$, où $\pi_{i++}, \pi_{+j+}, \pi_{++k}$ sont les termes marginaux à trois dimensions, on peut calculer la valeur de $\hat{\delta}^2$ explicitement en fonction des estimations des effets de conception des termes marginaux à trois dimensions et des estimations des effets de conception des cellules.

52. Les performances relatives de ces statistiques modifiées et de la statistique non modifiée sont indiquées au tableau XX.6 [de Rao et Scott (1984)] sur la base d'un tableau $2 \times 5 \times 4$ de l'Enquête de 1978-1979 sur la santé au Canada. Les variables sont le sexe ($I = 2$), l'utilisation de médicaments ($J = 5$) et le groupe d'âge ($K = 4$). Les hypothèses testées étaient : a) une complète indépendance (indiquée par $\bar{1} \otimes \bar{2} \otimes \bar{3}$); b) une indépendance partielle, par exemple $\pi_{ijk} = \pi_{i++}\pi_{+jk} \Leftrightarrow \bar{1} \otimes \bar{2} \bar{3}$

et, de même, $(\bar{2} \otimes \bar{1}\bar{3})$ et $(\bar{3} \otimes \bar{1}\bar{2})$; c) une indépendance conditionnelle (par exemple, $\pi_{ijk} = \pi_{i+k}\pi_{+jk} / \pi_{++k} \Leftrightarrow \bar{1} \otimes \bar{2}\bar{3}$) et, de même, $(\bar{2} \otimes \bar{1}\bar{3})$ et $(\bar{3} \otimes \bar{1}\bar{2})$. La conception était complexe, comportant une stratification et un échantillonnage à phases multiples. En outre, une post-stratification a été utilisée pour améliorer les estimations.

Tableau XX.6

Niveaux estimatifs de signification asymptotique (SL) de X^2 et des statistiques corrigées

$X^2/\hat{\delta}^2, X^2/\hat{\lambda}^2, X^2/\hat{d}^2$: tableau $2 \times 5 \times 4$ et niveau de signification nominale $\alpha = 0,05$

	Hypothèse						
	(a)		(b)			(c)	
	$\bar{1} \otimes \bar{2} \otimes \bar{3}$	$\bar{1} \otimes \bar{2}\bar{3}$	$(\bar{2} \otimes \bar{1}\bar{3})$	$(\bar{3} \otimes \bar{1}\bar{2})$	$\bar{1} \otimes \bar{2}\bar{3}$	$\bar{2} \otimes \bar{1}\bar{3}$	$\bar{3} \otimes \bar{1}\bar{2}$
SL (X^2)	0,72	0,33	0,76	0,72	0,43	0,30	0,78
SL ($X^2/\hat{\delta}^2$)	0,16	0,11	0,14	0,13	0,095	0,11	0,12
SL ($X^2/\hat{\lambda}^2$)	0,34	0,056	0,39	0,32	0,098	0,06	0,39
SL (X^2/\hat{d}^2)	0,34	0,054	0,39	0,32	0,097	0,06	0,39
$\hat{\delta}^2$	2,09	1,40	2,25	2,09	1,63	1,39	2,31
C.V. $\hat{\delta}_i$	1,54	1,02	1,37	1,27	0,86	1,05	1,11

53. Les comparaisons se font entre les niveaux de signification effectifs (SL) et le niveau nominal désiré, $\alpha = 0,05$. Une fois encore, les résultats affichent des valeurs inacceptables de SL (trop élevées) pour la statistique non corrigée. Les statistiques modifiées $X^2/\hat{\lambda}^2$ et X^2/\hat{d}^2 , qui ne dépendent pas de l'hypothèse, se comportent de façon très similaire, avec des valeurs de SL allant de 0,06 à 0,39, qui sont trop élevées. La modification fondée sur les effets marginaux et sur les effets de conception des cellules, $X^2/\hat{\delta}^2$, donne des résultats plus stables, avec des valeurs de SL allant de 0,095 à 0,16, toutes supérieures au niveau nominal, probablement en raison du fort coefficient de variation (CV) du terme $\hat{\delta}_i^2$.

54. En résumé, on dispose de méthodes de correction pour les statistiques de test chi-carré standard dans l'analyse des données catégoriques. Ces corrections sont souvent nécessaires pour permettre une analyse valable, avec un échantillon en grappe, et peuvent être faites assez facilement à l'aide des effets marginaux estimatifs et des effets de conception des cellules. On trouvera plus de détails au chapitre XXI sur les logiciels à utiliser pour faire face aux effets d'un échantillon de conception complexe sur les tests chi-carré et la régression logistique.

E. Résumé et conclusions

55. Dans ce chapitre, nous avons présenté des méthodes d'évaluation des effets d'échantillons de conception complexe couramment utilisés sur l'analyse des données d'enquêtes. Il s'agissait surtout de donner un premier aperçu des questions plutôt que d'offrir des prescriptions. L'évaluation et le traitement des effets de conception des échantillons sur l'analyse peuvent être difficiles et ne se prêtent pas à la formulation de règles faciles à appliquer. Comme nous l'avons montré, différents problèmes peuvent se prêter à différentes solutions, qui dépendent pour beaucoup du modèle proposé et de la validité des hypothèses sur lesquelles il repose, de divers aspects de la conception

de l'échantillon (par exemple, probabilités de sélection inégales, mise en grappe, etc.) et du type d'analyse envisagé. Il est essentiel de connaître la relation qui existe entre le modèle et les variables de conception de l'échantillon. Malheureusement, ces informations ne sont pas toujours immédiatement disponibles, ce qui veut dire que l'on peut avoir à utiliser des hypothèses et des approximations à leur place.

56. La condition première et fondamentale de toute analyse est la définition correcte du modèle sous-jacent. Cela incombe à l'analyste, mais l'identification finale du modèle peut et doit se fonder sur des techniques statistiques appropriées. L'analyse exploratoire initiale nécessaire pour identifier le modèle approprié peut se faire selon des méthodes graphiques et descriptives standard, sans que l'on ait à tenir compte des effets de la conception de l'échantillon.

57. Une fois qu'un modèle initial acceptable a été formulé, il faut déterminer si sa conception risque de rendre l'analyse confuse. Cela peut se faire, par exemple, à l'aide d'un test comparant les estimations pondérées et non pondérées des coefficients de régression linéaire (voir section C.2). S'il faut incorporer la conception complexe à l'analyse, il faut alors choisir la bonne façon de procéder. L'approche désagrégée ajoute simplement des variables au modèle lié à la conception de l'échantillon.

58. Souvent, toutefois, on ne peut modifier le modèle de manière à refléter pleinement les effets de la conception de l'échantillon de façon significative. Lorsque tel est le cas et qu'il faut utiliser l'approche agrégée, deux méthodologies fondamentales permettent de traiter de l'impact potentiel de cette conception. La première implique la modification des outils analytiques classiques pour tenir compte de la conception. C'est la méthode qui convient le mieux pour l'analyse de données catégoriques, où les statistiques chi-carré standard peuvent être modifiées sur la base d'effets de conception généralisés. La seconde approche consiste à mettre au point des outils analytiques spécialement conçus pour la conception. Des estimateurs pondérés en fonction de l'échantillon et une statistique de Wald fondée sur un vaste échantillon ont été proposés. Avant de pouvoir utiliser la statistique de Wald, il faut un estimateur fiable de la matrice de covariance, dont on ne dispose pas toujours dans la pratique.

59. Des travaux de recherche considérables sur les problèmes liés à la façon de traiter les effets d'un échantillon de conception complexe sur l'analyse ont permis de définir des méthodes pratiques, dont certaines ont été décrites dans le présent chapitre. D'autres travaux de recherche sont en cours, et nombre des méthodes existantes ont déjà été incorporées dans les logiciels nouveaux et existants. Malheureusement, du fait de la complexité du problème, il y a peu de chances qu'une méthode universelle voie le jour à l'avenir. Les méthodes et logiciels existants sont à utiliser avec une extrême prudence. Leur application nécessite à la fois une bonne connaissance de la théorie et une compréhension approfondie et une expérience de la construction de modèles pratiques.

Annexe

Définitions formelles et résultats techniques

Modèles de régression (sect. B.2 et B.3)

- Modèle de régression linéaire standard : $Y_i = \alpha + \beta X_i + \varepsilon_i$, où $\varepsilon_i | X_i \sim N(0, \sigma^2)$

- Coefficient de régression standard de la population : $B = \frac{\sum_{i=1}^N Y_i (X_i - \bar{X})}{\sum_{i=1}^N (X_i - \bar{X})^2}$

- Estimateur des moindres carrés ordinaires (OLS) de β : $b = \frac{\sum_{i=1}^n y_i (x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2}$

- Estimateur non biaisé de la variance de b : $v(b) = \frac{s^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$,

où s^2 est l'estimateur non biaisé de σ^2 , fondé sur la variance des termes résiduels de la régression.

- Modèle général de régression linéaire (hétéroscédastique) : $Y_i = \alpha + \beta X_i + \varepsilon_i$,

avec $\varepsilon_i | X_i \sim N(0, \sigma_i^2)$

- Coefficient de régression pondéré de la population : $B^* = \frac{\sum_{i=1}^N Y_i (X_i - \bar{X}_\sigma) / \sigma_i^2}{\sum_{i=1}^N (X_i - \bar{X}_\sigma)^2 / \sigma_i^2}$,

où $\bar{X}_\sigma = \frac{\sum_{i=1}^N X_i / \sigma_i^2}{\sum_{i=1}^N 1 / \sigma_i^2}$

- Estimateur généralisé des moindres carrés (GLS) de β : $b_G = \frac{\sum_{i=1}^n y_i (x_i - \bar{x}_\sigma) / \sigma_i^2}{\sum_{i=1}^n (x_i - \bar{x}_\sigma)^2 / \sigma_i^2}$,

où $\bar{x}_\sigma = \frac{\sum_{i=1}^n x_i / \sigma_i^2}{\sum_{i=1}^n 1 / \sigma_i^2}$

- Variance de l'estimateur GLS : $v(b_G) = \frac{1}{\sum_{i=1}^n (x_i - \bar{x}_\sigma)^2 / \sigma_i^2}$

- Estimateur de conception pondéré (Horvitz-Thompson) : $b^* = \frac{\sum_{i=1}^n w_i y_i (x_i - \bar{x}^*)}{\sum_{i=1}^n w_i (x_i - \bar{x}^*)^2}$,

où $w_i = \frac{1/\pi_i}{\sum_{k=1}^n 1/\pi_k}$, π_i est la probabilité d'inclusion, et $\bar{x}^* = \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i}$

Effet de l'exclusion des variables de conception (sect. C.1)

- Modèle étudié : $E(X_1) = \mu_1 + \beta_{12}(X_2 - \mu_2)$, où β_{12} est le paramètre étudié
- Modèle complet avec variable de conception : X_3

$$E(X_1) = \mu_1 + \beta_{12.3}(X_2 - \mu_2) + \beta_{13.2}(X_3 - \mu_3)$$

- Notation :

- Notation habituelle pour l'analyse multivariée, par exemple, $\beta_{12.3}$ est le coefficient de régression conditionnelle de X_1 sur X_2 , étant donné X_3

- Premier et deuxième moments X_i de la population : $\bar{X}_i = \frac{1}{N} \sum_{j=1}^N X_{ij}$

$$S_i^2 = \frac{1}{N-1} \sum_{j=1}^N (X_{ij} - \bar{X}_i)^2;$$

$$S_{ik} = \frac{1}{N-1} \sum_{j=1}^N (X_{ij} - \bar{X}_i)(X_{kj} - \bar{X}_k)$$

- Moments d'échantillon : $\bar{x}_i = \frac{1}{n} \sum_{j=1}^n x_{ij}$;

$s_{ik} = \frac{1}{n-1} \sum_{j=1}^n (x_{ij} - \bar{x}_i)(x_{kj} - \bar{x}_k)$; $s_i^2 = s_{ii}$, où nous supposons un échantillon, S , de taille fixe n , sélectionné selon une conception quelconque, dépendant éventuellement de X_3 .

- Estimateur OLS standard OLS de β_{12} : $b_{12} = \frac{S_{12}}{S_2^2}$

- Modèle asymptotique d'attente conditionnelle de b_{12} :

$$E_M(b_{12} | X_3, S) = \frac{\beta_{12} + \beta_{13}\beta_{23}(s_3^2/\sigma_3^2 - 1)}{1 + \rho_{23}^2(s_3^2/\sigma_3^2 - 1)} + O(n^{-1})$$

- Attente inconditionnelle (modèle et conception) :

$$E_M(b_{12}) = \beta_{12} + \frac{\sigma_1}{\sigma_2} \frac{\rho_{13.2}\rho_{23} \left[\left((1 - \rho_{12}^2)(1 - \rho_{23}^2) \right)^{\frac{1}{2}} (Q - 1) \right]}{1 + \rho_{23}^2(Q - 1)} + O(n^{-1}), \text{ où } Q = E(s_3^2)/\sigma_3^2$$

- L'estimateur OLS, b_{12} , est asymptotiquement biaisé, même inconditionnellement, à moins que $\rho_{23} = 0$ ou $E(s_3^2) = \sigma_3^2$, c'est-à-dire, $Q = 1$

- Estimateur non biaisé corrigé asymptotiquement [estimateur de vraisemblance maximum (MLE) dans des conditions normales] :

$$\hat{\beta}_{12} = \frac{s_{12} + (s_{13}s_{23}/s_3^2)(S_3^2/s_3^2 - 1)}{s_2^2 + (s_{23}^2/s_3^2)(S_3^2/s_3^2 - 1)}$$

- Estimateurs pondérés : $b_{12}^* = \frac{s_{12}^*}{s_2^{*2}}$; $\hat{\beta}_{12}^* = \frac{s_{12}^* + (s_{13}^*s_{23}^*/s_3^{*2})(S_3^{*2}/s_3^{*2} - 1)}{s_2^{*2} + (s_{23}^{*2}/s_3^{*2})(S_3^{*2}/s_3^{*2} - 1)}$, où

$\bar{x}_i = \sum_{j=1}^n \frac{x_{ij}}{N\pi_j}$; $s_{ik}^* = \sum_{j=1}^n \frac{x_{ij}x_{kj}}{N\pi_j} - \bar{x}_i\bar{x}_k$; $s_i^{*2} = s_{ii}^*$, et $\pi_j = p(j \in S | X_{3j})$ sont les probabilités d'inclusion dans l'échantillon.

Noter que $\sum_{j=1}^n \frac{1}{N\pi_j} = 1$ en cas d'échantillonnage aléatoire simple stratifié, qui est la conception que nous supposons ici. Pour des conceptions plus générales, $N\pi_j$ peut être remplacé par $1/w_j$, où :

$$w_j = \frac{1/\pi_j}{\sum_{j=1}^n 1/\pi_k}$$

- Résultat : $E_p(b_{12}^*) = E_p(\hat{\beta}_{12}^*) = B_{12} + O(n^{-1})$, où E_p est l'attente de conception (c'est-à-dire l'attente en cas de sélection répétée d'échantillons).

Analyse de données catégoriques (sect. D)

- Évaluation de la bonne adaptation :
 - Supposons une distribution multinomiale connue avec des probabilités $P_0 = (p_{0,1}, \dots, p_{0,k-1})$, où k est le nombre de catégories et $\sum_1^k p_{0,i} = 1$.
 - Avec H_0 , la statistique de chi-carré $X^2 = n \sum_{i=1}^k \frac{(\hat{p}_i - p_{0i})^2}{p_{0i}}$

(où \hat{p}_i sont des estimations d'échantillon de P_{0i}) est distribuée de façon asymptotique comme :

$X^2 = \sum_{i=1}^{k-1} \lambda_{0i}^2 Z_i^2$; $Z_i \underset{ind.}{\sim} N(0,1)$, où λ_{0i}^2 sont les valeurs *eigen* de $D = P_0^{-1}V_0$; P_0 est la matrice de variance des estimations de l'échantillon, selon l'hypothèse nulle pour SRS, et V_0 est leur matrice de vraie variance selon H_0 .

- Statistique de chi-carré modifiée :

$$X_C^2 = X^2 / \hat{\lambda}^2; \quad \hat{\lambda}^2 = \sum_{i=1}^{k-1} (1 - \hat{p}_i) \hat{d}_i^2 / (k-1)$$

où \hat{d}_i^2 sont les estimations des effets de conception, d_i^2 , de \hat{p}_i .

- Test d'indépendance dans les tableaux de contingences à deux dimensions :
 - Hypothèse étudiée : $H_0 : h_{ij} = p_{ij} - p_{i+}p_{+j} = 0$ ($i = 1, \dots, r; j = 1, \dots, c$), où p_{ij} est la proportion de la population de la $(i,j)^e$ cellule, et $p_{i+} = \sum_1^c p_{ij}$, $p_{+j} = \sum_1^r p_{ij}$ sont les totaux marginaux.
 - Statistique de chi-carré habituelle :

$$X_I^2 = n \sum_{j=1}^c \sum_{i=1}^r \frac{(\hat{p}_{ij} - p_{i+}p_{+j})^2}{p_{i+}p_{+j}}$$

où \hat{p}_{ij} est l'estimateur de l'échantillon de p_{ij} .

- X_I^2 est distribué de façon asymptotique comme la somme pondérée de b variables aléatoires indépendantes χ_1^2 .
- Correction de premier ordre : $X_{I(C)}^2 = X_I^2 / \hat{\delta}^2$.

où : $\hat{\delta}^2 = \sum_{i=1}^r \sum_{j=1}^c (1 - \hat{p}_{i+})(1 - \hat{p}_{+j}) \hat{\delta}_{ji}^2 / b$, et $\hat{\delta}_{ji}^2 = n \frac{\hat{v}(\hat{h}_{ij})}{\hat{p}_{i+} \hat{p}_{+j} (1 - \hat{p}_{i+})(1 - \hat{p}_{+j})}$ est l'estimation des effets de conception de \hat{h}_{ij} .

- On obtient une autre modification en remplaçant $\hat{\delta}^2$ par :

$$\hat{\lambda}^2 = \frac{1}{rc - 1} \sum_{i=1}^r \sum_{j=1}^c (1 - \hat{p}_{ij}) \hat{d}_{ij}$$

- Généralisations pour les modèles log-linéaires

- Modèle log-linéaire : $\mu = \tilde{\mu}(\theta)1 + X\theta$, où π est le vecteur- T des proportions des cellules de population, π_i , μ est le vecteur- T des log-probabilités $\mu_i = \ln \pi_i$, X est une matrice $T \times r$ connue de plein rang et $X'1 = 0$; θ est un vecteur- r des paramètres et $\tilde{\mu}(\theta) = 1/\{1/\exp(X\theta)\}$ est un facteur de normalisation.
- Hypothèse étudiée : $H_0 : \theta_2 = 0$, où $X = (X_1, X_2)$ et $\theta = (\theta_1, \theta_2)$, X_1 est $T \times s$ et X_2 est $T \times u$; θ_1 est $s \times 1$ et θ_2 est $u \times 1$.
- Soit $\hat{\pi}$ le pseudo-estimateur de vraisemblance maximum de π , avec M_1 , la solution des équations de pseudo-vraisemblance : $X'\hat{\pi} = X'\hat{p}$, où \hat{p} est un estimateur persistant (de conception) de π , selon la conception de l'enquête. De même, soit $\hat{\pi}$ le pseudo-estimateur de vraisemblance maximum de π , avec M_2 .

- Statistique standard de chi-carré de Pearson pour tester H_0 :

$$X^2 = n \sum_t \frac{(\hat{\pi}_t - \pi_t)^2}{\hat{\pi}_t}$$

- Distribution asymptotique de X^2 : $X^2 = \sum_{i=1}^u \delta_i^2 Z_i^2$; $Z_i \sim N(0,1)$ ind.

où δ_i^2 sont les valeurs *eigen* d'une matrice généralisée des effets de conception.

- Statistique de chi-carré modifiée : où : $X^2/\check{\delta}^2$, X^2/\check{d}^2 et X^2/\check{d}^2

- $\check{\delta}^2$ est l'estimation de la moyenne des valeurs *eigen*, $\check{\delta}^2 = \frac{1}{u} \sum_i \delta_i^2$

- $\hat{\lambda}^2$ est l'estimation de la moyenne des effets de conception de $X' \hat{p}$

- $\hat{d}^2 = \frac{1}{T} \sum_t \hat{d}_t^2$, où $\hat{d}_t^2 = n \frac{v(p_t)}{\pi_t(1-\pi_t)}$ est l'estimation de l'effet de conception de la cellule t .

- Exemple : pour l'hypothèse de l'indépendance complète dans un tableau à trois dimensions $I \times J \times K$, $H_0 : \pi_{ijk} = \pi_{i++} \pi_{+j+} \pi_{++k}$, où π_{i++} , π_{+j+} , π_{++k} sont les termes marginaux du tableau à trois dimensions; la valeur de $\check{\delta}^2$ est donnée par :

$$\check{\delta}^2 = \frac{\sum_i \sum_j \sum_k (1 - \hat{\pi}_{i++} \hat{\pi}_{+j+} \hat{\pi}_{++k}) \hat{d}_{ijk}^2 - \sum_i (1 - \hat{\pi}_{i++}) \hat{d}_i^2(r) - \sum_j (1 - \hat{\pi}_{+j+}) \hat{d}_j^2(c) - \sum_k (1 - \hat{\pi}_{++k}) \hat{d}_k^2(l)}{IJK - I - J - K + 2}$$

où $\hat{d}_i^2(r)$, $\hat{d}_j^2(c)$, et $\hat{d}_k^2(l)$ sont des estimations des effets de conception des termes marginaux du tableau à trois dimensions et \hat{d}_{ijk}^2 est l'estimation de l'effet de conception des cellules.

RÉFÉRENCES

- An, A., et D. Watts (2001). *New SAS Procedures for Analysis of Sample Survey Data*. SAS Users Group International (SUGI) paper, No. 23. Cary, Caroline du Nord, SAS Institute, Inc. Disponible sur <http://www2.sas.com/proceedings/sugi23/stats/p247.pdf> (consulté le 2 juillet 2004).
- Berthoud, R., et J. Gershuny, eds. (2000). *Seven Years in the Lives of British Families: Evidence on the Dynamics of Social Change from the British Household Panel Survey*. Bristol, Royaume-Uni, The Policy Press.

- Brewer, K. R. W., et R. W. Mellor (1973). The effect of sample structure on analytical surveys. *Australian Journal of Statistics*, vol. 15, pp. 145-152.
- Chambers, R. L., et C. J. Skinner, eds. (2003). *Analysis of Survey Data*. New York, Wiley and Sons, Inc.
- Cochran, W. G. (1977). *Sampling Techniques*, 3^e éd. New York, Wiley and Sons, Inc.
- DuMouchel, W. H., et G. J. Duncan (1983). Using sample survey weights in multiple regression analyses of stratified samples. *Journal of the American Statistical Association*, vol. 78, pp. 535-543.
- Duncan, G. J., et G. Kalton (1987). Issues of design and analysis of surveys across time. *International Statistical Review*, vol. 55, pp. 97-117.
- Feder, M., G. Nathan et D. Pfeffermann (2000). Time series multi-level modelling of complex survey longitudinal data with time varying random effects. *Survey Methodology*, vol. 26, pp. 53-65.
- Fuller, W. A. (1984). Least squares and related analyses for complex survey designs. *Survey Methodology*, vol. 10, pp. 97-118.
- Hansen, M. H., W. G. Madow et B. J. Tepping (1983). An evaluation of model-dependent and probability-sampling inferences in sample surveys. *Journal of the American Statistical Association*, vol. 78, pp. 776-793.
- Kish, L., et M. Frankel (1974). Inference from complex samples. *Journal of the Royal Statistical Society B*, vol. 36, pp. 1-37.
- Korn, E. L., et B. I. Graubard (1999). *Analysis of Health Surveys*. New York et Chichester, Royaume-Uni, Wiley and Sons, Inc.
- Manski, C. F., et S. R. Lerman (1977). The estimation of choice probabilities from choice based samples. *Econometrica*, vol. 45, pp. 1977-1988.
- Nathan, G. (1972). On the asymptotic power of tests of independence in contingency tables from stratified samples. *Journal of the American Statistical Association*, vol. 67, pp. 917- 920.
- _____, et D. Holt (1980). The effect of survey design on regression analysis. *Journal of the Royal Statistical Society B*, vol. 43, pp. 377-386.
- _____, et T. M. F. Smith (1989). The effect of selection on regression analysis. In *Analysis of Complex Surveys*, C. J. Skinner, D. Holt et T. M. F. Smith, eds. Chichester, Royaume-Uni, Wiley et Sons, Inc., pp. 227-250.
- Pfeffermann, D. (1993). The role of sampling weights when modeling survey data. *International Statistical Review*, vol. 61, pp. 317-337.
- _____, et Sverchkov, M. (1999). Parametric and semi-parametric estimation of regression models fitted to survey data. *Sankhya, Series B*, vol. 61, part. 1, pp. 166-186.
- _____, F. Moura et N. S. Silva (2001). Multi-level modelling under informative probability sampling. Invited paper for the 53rd Session of the International Statistical Institute, Séoul.
- _____, et al. (1998). Weighting for unequal selection probabilities in multi-level models. *Journal of the Royal Statistical Society B*, vol. 60, pp. 23-40.
- Rao, J.N.K., et A. J. Scott (1981). The analysis of categorical data from complex sample surveys: hi-squared tests for of fit and independence in two-way tables. *Journal of the American Statistical Association*, vol. 76, pp. 221-230.
- _____, (1984). On chi-squared tests for multi-way contingency tables with cell proportions estimated from survey data. *Annals of Statistics*, vol. 12, pp. 46-60.
- Särndal, C.-E., B. Swensson et J. Wretman (1992). *Model Assisted Survey Sampling*. New York, Springer-Verlag.

Scott, A. J., et C. J. Wild (1986). Fitting logistic models under case control or choice-based sampling. *Journal of the Royal Statistical Society B*, vol. 48, pp. 170-182.

Skinner, C. J., D. Holt et T. M. F. Smith, eds. (1989). *Analysis of Complex Surveys*. Chichester, Royaume-Uni, Wiley and Sons, Inc.

Valliant, R., A. H. Dorfman et R. M. Royall (2000). *Finite Population Sampling and Inference: A Prediction Approach*. Chichester, Royaume-Uni, et New York, Wiley and Sons, Inc.

Chapitre XXI

Estimation de l'erreur d'échantillonnage pour les données d'enquête*

DONNA BROGAN
 Université Emory
 Atlanta, Géorgie
 États-Unis d'Amérique

RÉSUMÉ

Les échantillons d'enquête de conception complexe s'écartent de l'échantillonnage aléatoire simple, y compris pour certains aspects tels que l'échantillonnage probabiliste inégal, l'échantillonnage à phases multiples et la stratification. On a besoin d'analyses pondérées pour établir des estimations non biaisées (ou quasi non biaisées) des paramètres de population. L'estimation de la variance des estimateurs dépend des caractéristiques du plan d'échantillonnage et nécessite des méthodes d'approximation, généralement la linéarisation (série de Taylor) ou des techniques de réplification.

Généralement, on ne peut utiliser les logiciels de statistique courants pour analyser les données d'une enquête par sondage, car ces logiciels supposent le plus souvent des éléments d'échantillonnage aléatoire simple. Ces logiciels produisent des estimations ponctuelles biaisées des paramètres de la population (dans une analyse non pondérée) et/ou une sous-estimation des erreurs types de ces estimations. L'utilisation de la variable d'échantillonnage pondéré avec des logiciels courants produit des estimations ponctuelles appropriées des paramètres de population. Toutefois, les estimations des erreurs types restent généralement inexactes parce que la procédure d'estimation de la variance ne tient généralement pas compte de la mise en grappe et/ou de la stratification du plan d'échantillonnage.

Le présent chapitre donne un aperçu de huit logiciels offrant la capacité d'analyse des données d'enquêtes par sondage, y compris le coût approximatif, les méthodes d'estimation de la variance, les options en matière d'analyse, l'interface avec l'utilisateur et les avantages et inconvénients. Quatre de ces logiciels sont gratuits, et peuvent donc intéresser les pays en développement qui ne disposent que d'un budget limité pour l'acquisition de logiciels.

Un ensemble complexe de données d'enquête par sondage du Burundi illustre les inexactitudes obtenues par les analyses effectuées à l'aide d'un logiciel de statistique courant. Des analyses descriptives annotées de l'enquête effectuée au Burundi pour cinq des huit logiciels examinés (STATA, SAS, SUDAAN, WesVar et Epi-Info) montrent comment utiliser ces logiciels. Enfin, on compare à des fins analytiques les résultats numériques obtenus à l'aide de ces cinq logiciels aux données de l'enquête effectuée au Burundi. Tous les cinq logiciels donnent une estimation équivalente de la variance, que la méthode utilisée soit la linéarisation (série de Taylor) ou la réplification équilibrée répétée [*balanced repeated replication (BRR)*].

* Ce chapitre comprend une Annexe (en anglais seulement) contenant des analyses présentées aux fins d'illustration et de comparaison des données de l'Enquête sur l'immunisation au Burundi, effectuées à l'aide de cinq logiciels de statistiques. Le contenu du CD-ROM, y compris les codes de programme et les résultats pour chacun de ces logiciels, peut être téléchargé directement du site Internet de la Division de la statistique des Nations Unies (<http://unstats.un.org/unsd/hhsurveys/>) ou le CD-ROM peut être obtenu sur demande adressée à la Division de la statistique des Nations Unies (statistics@un.org).

Termes clés : linéarisation par série de Taylor, méthodes de réplification, grappe ultime, estimation de la variance, enquêtes par sondage complexes, logiciels.

A. Conceptions d'échantillons d'enquête

1. Comme l'ont montré de nombreux chapitres de cette publication, les échantillons conçus pour les enquêtes sur les ménages sont complexes et comportent généralement un échantillonnage stratifié à phases multiples. L'utilisation d'un échantillon de conception complexe a notamment pour conséquence que les méthodes et les logiciels statistiques courants ne peuvent s'appliquer sans discernement à l'analyse des données de l'enquête. En particulier, on affecte aux unités qui répondent à une enquête des pondérations qui compensent l'inégalité de probabilité de sélection, les cas de non-réponse et de non-couverture et qui peuvent être utilisées pour faire en sorte que les distributions pondérées de certaines variables concordent avec les distributions connues de ces variables. Ces pondérations doivent être utilisées dans l'analyse de l'enquête. En outre, le calcul des erreurs d'échantillonnage pour les estimations de l'enquête doit tenir compte du fait que l'on a choisi l'échantillon de l'enquête en recourant à une conception complexe de l'échantillon. Heureusement, il existe à présent un certain nombre de logiciels spécialisés pour l'analyse d'une enquête qui calculent correctement les erreurs d'échantillonnage pour les estimations pondérées provenant d'échantillons de conception complexe. Nous présenterons et examinerons certains de ces logiciels dans le présent chapitre.

2. À titre de préalable à l'examen des logiciels d'analyse des enquêtes, les deux sections suivantes traitent de la question des analyses pondérées et des méthodes d'estimation de la variance pour les échantillons de conception complexe. Les sections ultérieures sont consacrées à une comparaison des huit logiciels d'estimation de la variance pour les estimations tirées des données d'une enquête par sondage complexe et décrivent l'utilisation de cinq d'entre eux sur les données d'une enquête par sondage effectuée au Burundi. Le chapitre se termine par certaines conclusions et recommandations. L'annexe figurant sur le CD-ROM produit parallèlement à cette publication offre des analyses annotées des données pour trois analyses effectuées à l'aide des cinq logiciels choisis (voir note en début de chapitre).

B. Questions liées à l'analyse des données d'enquêtes par sondage complexes

1. Analyses pondérées

3. Dans beaucoup d'enquêtes sur les ménages, les unités analysées – ménages ou personnes – sont choisies selon des probabilités inégales, et il faut des pondérations pour compenser ces inégalités dans les analyses. De plus, même lorsque les unités sont choisies avec une égale probabilité, il faut souvent des pondérations pour compenser les cas de non-réponse unitaire et pour l'analyse comparative, par exemple la post-stratification (voir chapitre XIX). Ces pondérations doivent être utilisées dans les analyses pour estimer les paramètres de population. Selon l'enquête, les estimateurs non pondérés (non recommandés) peuvent être fortement biaisés pour les paramètres de population. La valeur de la variable de pondération de l'échantillon, appelée *WTVAR*, pour un élément donné *R* d'un échantillon de répondants dans l'ensemble de données peut être interprété comme le nombre d'éléments de la population représenté par *R*. La somme des valeurs de *WTVAR* pour tous les éléments *R* des estimations de l'ensemble de données est une estimation du nombre d'éléments de la population.

4. Parfois, on « normalise » la variable $WTVAR$ de pondération de l'échantillon en la multipliant par (nombre de R) / (somme des valeurs de $WTVAR$ de tous les R). La somme de la valeur de la « variable de pondération normalisée » $WTNORM$ de tous les R est la taille de l'échantillon de l'analyse (nombre de R). Il n'importe pas que l'on utilise la variable de pondération de l'échantillon $WTVAR$ ou la variable de pondération normalisée $WTNORM$ pour obtenir une estimation ponctuelle d'un paramètre « moyen » de population tel qu'une moyenne ou une proportion : les deux donnent le même calcul. En revanche, la variable de pondération normalisée $WTNORM$ ne peut être utilisée pour estimer directement les totaux de certains paramètres de population, tels que le nombre total d'enfants souffrant de malnutrition au sein d'une population.

2. Aperçu de l'estimation de la variance

5. L'estimation de la variance est importante parce qu'elle indique la précision des estimateurs et permet d'établir des intervalles de confiance pour les paramètres de population et de tester les hypothèses concernant ces paramètres. L'estimation de la variance pour des estimateurs fondés sur les données d'une enquête par sondage complexe doivent tenir compte des facteurs suivants : *a*) la plupart des estimateurs sont non linéaires (un ratio d'estimateurs linéaires est chose courante); *b*) les estimateurs sont pondérés; *c*) le plan d'échantillonnage a généralement utilisé la stratification avant l'échantillonnage de première phase (et parfois également aux phases d'échantillonnage suivantes); et *d*) généralement, du fait de l'échantillonnage en grappe à phases multiples, les éléments de l'échantillon ne sont pas statistiquement indépendants. Dans presque tous les cas, on ne peut obtenir d'expression algébrique fermée de l'estimation de la variance. Aussi les ouvrages de recherche sur l'estimation de la variance de données d'enquêtes par sondage complexes présentent-ils plusieurs méthodes d'approximation au choix des analystes des données de ces enquêtes.

6. Les deux méthodes d'approximation de l'estimation de la variance les plus couramment utilisées sont la linéarisation par série de Taylor [*Taylor series linearization* (TSL)] (Wolter, 1985; Shah, 1998) et les techniques de réplification (Wolter, 1985; Rust et Rao, 1996). Ces approches sont examinées de façon plus approfondie à la section C. La plupart des logiciels d'analyse des données d'enquêtes par sondage n'appliquent que l'une de ces deux méthodes. Pour les estimateurs qui sont immédiatement fonction des données de l'échantillon (par exemple, les totaux, les moyennes, les proportions, les différences entre moyennes/proportions, etc.), les deux méthodes donnent des estimations comparables de la variance et aucune des deux n'a clairement la préférence. Pour les estimateurs qui ne sont pas immédiatement fonction de ces données (par exemple, les valeurs médianes), une procédure particulière de réplification, la réplification équilibrée répétée, semble avoir la préférence sur la linéarisation par série de Taylor et « jackknife », autre méthode de réplification (Korn et Graubard, 1999). De nombreux ouvrages comparent les techniques d'estimation de la variance, y compris certains cas particuliers où une méthode est jugée préférable à une autre [par exemple, voir Korn et Graubard (1999) et leurs nombreuses références, et Kish et Frankel (1974)].

3. Facteurs de correction d'une population finie (*fpc*) pour l'échantillonnage sans remplacement

7. Pour plus de simplicité, considérons d'abord l'estimation d'une moyenne de population obtenue à partir d'un échantillon de taille n sélectionné selon une proba-

bilité égale à partir d'une population N , et comparons deux conceptions d'échantillon. Selon une conception, les éléments sont sélectionnés par échantillonnage aléatoire simple, c'est-à-dire sans remplacement. Selon l'autre conception, ils sont sélectionnés par échantillonnage sans restriction, c'est-à-dire avec remplacement (cette méthode est également dénommée échantillonnage aléatoire simple avec remplacement). La différence de variance pour la moyenne de l'échantillon selon ces deux conceptions est qu'un terme de correction (fpc) de la population finie est inclus dans la variance avec l'échantillon aléatoire simple mais pas dans la variance avec l'échantillon sans restriction (voir chapitre VI). Le terme fpc est $(1 - f)$ où $f = n/N$ est la fraction d'échantillonnage. Ce terme est supérieur de 1,0 et reflète la réduction de la variance résultant de l'échantillonnage sans remplacement. Si la fraction d'échantillonnage f est petite, le terme fpc est proche de 1,0 et a un impact minime sur la variance. On peut donc l'ignorer en toute sécurité dans l'estimation de la variance. Autrement dit, l'échantillon sans remplacement peut être traité comme s'il avait été sélectionné avec remplacement. Généralement, une fraction d'échantillonnage allant jusqu'à 5 ou 10 % est jugée petite. En revanche, si f est grande, le fait de ne pas tenir compte du terme fpc lors de la sélection de l'échantillon sans remplacement entraîne une surestimation de la variance. Dans la conception d'un échantillon aléatoire stratifié avec différentes fractions d'échantillonnage dans différentes strates, le terme fpc peut être suffisamment faible pour être ignoré dans certaines strates mais pas dans d'autres.

8. La plupart des enquêtes sur les ménages sont fondées sur l'application d'échantillons de conception complexe à des populations très nombreuses. Les UPE sont généralement choisies selon une probabilité proportionnelle à la taille [*probability proportional to size* (PPS)] sans échantillonnage de remplacement, rendant la notion de « fraction d'échantillonnage » plus complexe. Cependant, le nombre d'UPE est souvent élevé et la fraction d'échantillonnage des UPE de chaque strate est relativement petite, ce qui donne une valeur proche de 1,0 pour tous les termes fpc de la première phase. De ce fait, l'analyse des données d'une enquête par sondage complexe utilise souvent comme approximation le cas où les UPE ont été sélectionnées avec remplacement dans chaque strate. Si l'on fait cette approximation lorsque certaines strates ont de grandes fractions d'échantillonnage de première phase, on risque d'obtenir une variance sensiblement surestimée. Cette surestimation est souvent acceptée en raison de la complexité d'estimation de la variance en l'absence d'approximation. Il convient de noter que, si l'échantillonnage se fait sans remplacement à la première phase d'échantillonnage d'une strate quelconque, cette strate n'implique pas d'approximation.

4. Pseudo-strates et pseudo-UPE

9. Pour les besoins de l'estimation de la variance, parfois, les strates et les UPE ne sont pas identifiées car elles ont été effectivement utilisées dans le plan d'échantillonnage. On peut apporter des modifications à la définition des strates et des UPE pour l'estimation de la variance de manière à ce que le plan d'échantillonnage utilisé corresponde à l'une des options offertes par le logiciel pour ce plan. Dans ce cas, les nouvelles variables des strates et d'UPE définies pour estimer la variance sont parfois appelées pseudo-strates et pseudo-UPE.

10. L'exemple courant est celui où un très grand nombre de strates sont définies avant l'échantillonnage de première phase avec une seule UPE sélectionnée à l'intérieur de chaque strate. On ne peut estimer la variance avec une seule UPE par strate, car on ne peut estimer la variabilité des UPE à l'intérieur de la strate. Dans ce cas, on fusionne ou combine deux strates en une pseudo-strate, et on obtient ainsi deux UPE pour cette pseudo-strate. La fusion de strates se fait de façon stratégique et non pas ar-

bitraire, sur la base de la connaissance des variables de stratification des UPE et de la méthode de sélection des UPE (Kish, 1965).

11. L'autre exemple est celui de la stratification implicite. Par exemple, un pays peut être stratifié Nord-Sud, avec des PSU définies par les villages. À l'intérieur de chaque strate, les UPE sont ordonnées en fonction de leur proximité géographique, après quoi on sélectionne un échantillon probabiliste de nombreuses UPE (disons 30) à l'intérieur de chaque strate en utilisant systématiquement une probabilité proportionnelle à la taille estimative [*probability proportional to estimated size* (PPES)] (Kish, 1965). La classification géographique des UPE selon leur taille à l'intérieur des strates, combinée à un échantillonnage systématique, entraîne une stratification géographique implicite des villages (UPE) à l'intérieur des strates Nord et Sud. Pour reconnaître la stratification implicite dans l'estimation de la variance, le plan d'échantillonnage est généralement décrit comme regroupant 15 pseudo-strates Nord et 15 pseudo-strates Sud, comprenant chacune deux UPE sélectionnées ou pseudo-UPE. Les deux premières UPE sélectionnées à partir du cadre d'échantillonnage Nord iraient dans la première pseudo-strate et les deux UPE suivantes dans la seconde, etc.

12. Korn et Graubard (1999) donne plusieurs autres exemples de formation de pseudo-strates et de pseudo-UPE pour l'estimation de la variance, par exemple pour réduire le nombre de réplicats et le volume de calculs. Par ailleurs, l'appendice D du *WesVar User's Guide* (2002) donne des conseils et des exemples de description de divers logiciels de plans d'échantillonnage pour l'estimation de la variance fondés sur des méthodes de réplication.

5. Approximation courante (WR) pour décrire les plans d'échantillonnage complexes

13. Généralement, les enquêtes par sondage complexes utilisent un échantillonnage en grappe à phases multiples. Par ailleurs, la stratification des unités de la seconde phase et des phases suivantes (à l'intérieur d'une UPE sélectionnée) peut avoir lieu avant que l'échantillonnage ne se fasse pour ces phases. Toutefois, les méthodes d'approximation de l'estimation de la variance couramment utilisées pour ces cas complexes n'ont pas besoin de tenir compte de toutes les phases d'échantillonnage et de stratification. Aux phases ultérieures, l'échantillonnage complexe est automatiquement couvert par l'approximation « avec remplacement » de la première phase examinée précédemment. En fait, peu de logiciels d'enquête par sondage peuvent inclure séparément toutes les phases de l'échantillonnage dans l'estimation de la variance en l'absence d'une approximation de première phase avec remplacement.

14. Il est très courant d'utiliser l'estimation de la variance ultime de grappes [*ultimate cluster variance estimate* (UCVE)] pour les échantillons de conception complexe, proposée d'abord par Hansen, Hurwitz et Madow (1953) et traitée également par Wolter (1985). Cette estimation peut être effectuée avec la linéarisation par série de Taylor ou avec une technique de réplication. L'approche UCVE traite les UPE comme si elles étaient sélectionnées avec remplacement dans les strates de la première phase. Dès lors, chaque R (élément de l'ensemble de données correspondant à un répondant de l'échantillon) n'a besoin d'être identifié que par la strate de la première phase et l'UPE (à l'intérieur de cette strate) à partir desquelles il a été sélectionné. On n'a pas besoin d'informations sur les phases de l'échantillonnage en dessous du niveau des UPE mais avant la phase à laquelle correspond l'élément pour l'estimation de la variance. De la sorte, la description du plan d'échantillonnage proprement dit se trouve simplifié à tel point qu'il ressemble à un échantillon stratifié en grappe à une phase, c'est-à-dire à un échantillon stratifié de grappes ultimes totalement recensées.

Cette approche fondée sur la grappe ultime donne une bonne approximation pour l'estimation de la variance, à condition que l'hypothèse de la première phase avec remplacement soit raisonnable. Cette approximation courante (UCVE) est parfois identifiée par les lettres *WR* (*with replacement*) dans les études sur les enquêtes par sondage, et ces lettres ont la signification indiquée ci-dessous.

15. Ainsi, lorsque le plan d'échantillonnage est décrit comme *WR*, on n'a besoin que de trois variables de conception de l'enquête pour estimer la variance :

- a) La variable de pondération de l'échantillon *WTVAR* (également nécessaire pour les estimations ponctuelles);
- b) La variable de stratification (ou variable de pseudo-stratification) *STRATVAR* utilisée avant l'échantillonnage de première phase (UPE);
- c) La variable de l'UPE (ou pseudo-UPE), *PSUVAR*.

16. Chaque répondant *R* de l'échantillon doit avoir une valeur pour chacune de ces trois variables dans le fichier de données de base. Par exemple, un répondant particulier *R* peut représenter 8 714 éléments de la population (*WTVAR* a la valeur de 8 714) et peut avoir été sélectionné à partir de la strate ou pseudo-strate #6 (*STRATVAR* a la valeur de 6) et de l'UPE ou pseudo-UPE #3 de la strate 6 (une valeur de 3 pour *PSUVAR*, à l'intérieur de *STRATVAR* = 6).

17. *WR* est la description unique ou par défaut du plan d'échantillonnage donnée par la plupart des logiciels ou procédures d'enquête par sondage. C'est le cas de la linéarisation par série de Taylor, de SUDAAN, SAS, STATA, Epi-Info, PC-CARP et CENVAR. *WR* est la description par défaut donnée avec BRR et « jackknife », dans WesVar et SUDAAN. Il est à noter que l'échantillonnage à une seule phase, tel que l'échantillonnage aléatoire simple ou l'échantillonnage aléatoire stratifié, est un cas particulier de l'échantillonnage à phases multiples où les UPE du cadre d'échantillonnage sont les éléments de population et chaque UPE ne comprend qu'un élément (autrement dit, il n'y a pas de mise en grappe d'éléments d'échantillon). Les logiciels qui ne donnent qu'une description *WR* au plan d'échantillonnage peuvent offrir la possibilité d'incorporer les termes *fpc* dans l'estimation de la variance lorsque l'on utilise l'échantillonnage d'éléments sans remplacement (par exemple, SAS, STATA, WesVar).

18. Quand on utilise *WR* pour donner une approximation du plan d'échantillonnage complexe, on risque de surestimer légèrement les variances. Néanmoins, les analystes des données d'enquêtes sont prêts à accepter un certain degré de surestimation en échange de la relative simplicité de l'approximation *WR*. Il convient toutefois de noter que cette surestimation peut être appréciable s'il y a plusieurs strates lorsque l'échantillonnage de première phase est sans remplacement et comprend de fortes fractions d'échantillonnage. Dans ce cas, il peut être jugé souhaitable d'utiliser un logiciel qui incorpore les facteurs *fpc* de la première phase.

6. Techniques d'estimation de la variance et variables de conception de l'enquête

19. Généralement, les ensembles de données d'enquêtes par sondage à publier sont déjà prêts pour l'estimation de la variance selon l'une des deux principales approches, la linéarisation par série de Taylor ou les techniques de réplification. À l'occasion, l'ensemble de données destiné à être publié est conçu de manière à pouvoir utiliser ces deux approches de l'estimation de la variance. Il importe que les variables de conception de l'échantillon à utiliser pour estimer la variance figurent dans l'ensemble de

données à publier, accompagnées de la documentation correspondante sur la façon dont ces variables sont définies et sur la façon de les utiliser.

20. Si l'on utilise la linéarisation par série de Taylor, il faut chercher trois variables de conception de l'enquête dans la documentation : la variable de pondération de l'échantillon *WTVAR*, la variable de stratification de première phase *STRATVAR*, et la variable de l'UPE *PSUVAR* (bien entendu, ces variables n'auront pas les noms utilisés ici). Si l'on utilise une méthode de réplification, il faut rechercher la variable de pondération de l'échantillon *WTVAR* et plusieurs variables de répliqués, souvent appelées par des noms tels que REPL01--REPL52 (pour les variables de pondération de 52 répliqués). On n'a pas besoin de connaître les variables *STRATVAR* ou *PSUVAR* si l'ensemble de données contient les variables de pondération des répliqués.

21. Les enquêteurs qui font leur enquête et préparent leur propre ensemble de données à analyser ont besoin d'inclure les variables de conception de l'enquête pertinentes et de leur affecter une valeur pour chaque élément de répondant (R) dans l'ensemble de données. L'ensemble minimum de variables nécessaires est le suivant : la variable de pondération de l'échantillon *WTVAR*, la variable de stratification (ou de pseudo-stratification) de première phase *STRATVAR*, et la variable de l'UPE (ou de la pseudo-UPE) *PSUVAR* de la strate. Ces trois variables de conception de l'enquête donnent une approximation *WR* du plan d'échantillonnage effectif et permettent l'utilisation directe de la linéarisation par série de Taylor ou le calcul personnel ou par logiciel des pondérations des répliqués pour l'utilisation des techniques de réplification pour estimer la variance. Si l'on désire incorporer les termes *fpc* et/ou des phases additionnelles d'échantillonnage ou de stratification à l'estimation de la variance, on a besoin de variables additionnelles de conception de l'enquête dans l'ensemble de données ainsi que d'un logiciel d'enquête par sondage offrant ces possibilités (par exemple, SUDAAN).

22. Il est malheureusement fréquent que l'on acquière un ensemble de données d'enquête par sondage ne contenant aucune des variables de conception ou des variables de pondération des répliqués. Si un échantillonnage probabiliste a été utilisé, il faut reconstruire les variables de conception de l'enquête *WTVAR* pour l'estimation et les variables *STRATVAR* et *PSUVAR* pour l'estimation de la variance. Parfois, on peut obtenir suffisamment de détails du plan d'échantillonnage en consultant la documentation ou en contactant le personnel d'échantillonnage pour reconstruire les variables de conception de l'enquête. Si l'on ne dispose que d'informations limitées, on peut avoir à se contenter d'approximations grossières. Par exemple, si l'on ne peut construire les probabilités de sélection, on peut raisonnablement supposer un échantillon d'éléments d'égale probabilité et utiliser simplement un ajustement post-stratification pour obtenir les valeurs de *WTVAR*. Si l'on ne peut déterminer exactement les UPE, on peut concevoir une UPE de remplacement (proxy) si l'on dispose de certaines coordonnées géographiques, mais il faut alors tenir compte des limitations de l'analyse des données si les variables de conception de l'échantillon sont imprécises.

7. Analyse des données d'une enquête par sondage complexe

23. L'analyse des données d'une enquête par sondage complexe pose un grand nombre de problèmes théoriques et pratiques, par-delà l'exécution d'une analyse pondérée et de l'estimation correcte des variances des estimateurs. Ces problèmes sont bien analysés et illustrés dans l'ouvrage récent de Korn et Graubard (1999), qui traite notamment de certains sujets tels que l'adaptation des modèles (par exemple, la régression logistique) aux données d'enquêtes par sondage, la bonne adaptation pour les modèles, l'estimation de la variance pour des subdivisions de la population, qui com-

bine des enquêtes multiples et la formation de pseudo-strates et de pseudo-UPE (voir également d'autres chapitres de la présente partie de cette publication).

C. Méthodes d'estimation de la variance

1. Linéarisation par série de Taylor pour l'estimation de la variance

24. Supposons un plan d'échantillonnage complexe avec stratification des UPE, échantillonnage à phases multiples et probabilité inégale de sélection des éléments. L'estimateur linéaire $\sum w_i y_i$, qui est une somme pondérée, estime le total de population pour la variable y , où w_i est la valeur de la variable de pondération de l'échantillon *WTVAR* pour un élément i de l'échantillon, y_i est la valeur de la variable y pour l'élément i de l'échantillon, et la somme Σ est l'ensemble des éléments de l'échantillon, $i = 1, 2, \dots, m$. Si y est une variable binaire codée 1 pour les hommes diabétiques et 0 autrement, dès lors, le total de population à estimer est le nombre total d'hommes diabétiques. On peut déterminer directement la variance estimée de $\sum w_i y_i$ dans l'hypothèse *WR* examinée ci-dessus.

25. À présent, soit x_i une variable binaire codée 1 pour les hommes et 0 pour les femmes. Dès lors, la prévalence estimée du diabète chez les hommes est donné par $[\sum w_i y_i] / [\sum w_i x_i]$, qui est un ratio de deux estimateurs linéaires (ou de deux sommes pondérées). Dans l'hypothèse *WR*, on peut déterminer directement la variance estimée de ce ratio. Même si l'on a utilisé un échantillonnage aléatoire simple au lieu de méthodes d'échantillonnage complexes, l'estimation de la variance de cette fonction non linéaire, un ratio, n'est pas directe et nécessite une méthode d'approximation.

26. On peut développer l'expression algébrique de l'estimateur non linéaire ci-dessus en une série de Taylor infinie centrée sur la valeur (estimative) attendue du numérateur et la valeur (estimative) attendue du dénominateur. On obtient alors une approximation algébrique de l'estimateur non linéaire en ne retenant que les principaux termes de la série de Taylor infinie, ce qui se traduit par une expression algébrique qui est à présent une fonction linéaire (et non plus non linéaire) des données de l'échantillon, ce qui veut dire que l'estimateur non linéaire du ratio a été « linéarisé ». On peut alors obtenir directement la variance estimée de la fonction linéarisée (y compris les termes pertinents de la covariance) selon l'hypothèse *WR*, de la même façon que l'on a obtenu la variance estimée de $\sum w_i y_i$. Dans ce processus, on estime la variance de la fonction linéarisée séparément à l'intérieur de chaque strate (car l'échantillonnage est indépendant à travers les strates) et l'on fait la somme des variances estimatives propres à chaque strate pour obtenir la variance de l'estimateur.

27. Lorsque l'on utilise l'approche de la linéarisation par série de Taylor, il suffit de dériver une unique formule d'estimation de la variance approximative et de la programmer non seulement pour chaque estimateur non linéaire mais aussi pour chaque plan d'échantillonnage où cet estimateur pourrait être utilisé (*WR* étant un tel plan d'échantillonnage). Cette caractéristique est considérée comme un inconvénient de l'approche de l'estimation de la variance par la linéarisation par série de Taylor. En effet, un certain logiciel qui analyse les données d'une enquête par sondage selon la linéarisation par série de Taylor risque de ne pas inclure la combinaison de l'estimateur particulier que l'on désire utiliser avec le plan d'échantillonnage effectif ou approximatif que l'on a utilisé.

28. Tous les logiciels qui utilisent la linéarisation par série de Taylor nécessitent que l'on spécifie les variables de conception *WTVAR*, *STRATVAR* et *PSUVAR*, au même titre que pour l'approximation du plan d'échantillonnage *WR*. On peut dispo-

ser de plans d'échantillonnage additionnels avec la linéarisation par série de Taylor, selon le logiciel, mais leur utilisation peut nécessiter d'autres variables de conception.

2. Méthode de réplification pour l'estimation de la variance

29. Bien que la méthode de réplification soit déjà connue en théorie depuis un certain temps, son utilisation pour l'estimation de la variance d'estimateurs d'enquêtes par sondage s'est développée avec l'avènement de la capacité de calcul ultrarapide. La méthode de réplification fait abondamment usage de l'informatique, mais elle est plus souple que la linéarisation par série de Taylor en ce qui concerne le nombre d'estimateurs différents dont elle permet le calcul des variances estimatives.

30. L'idée générale des méthodes de réplification est la suivante : premièrement, elle utilise l'échantillon tout entier, comme la méthode de linéarisation par série de Taylor, pour obtenir une estimation ponctuelle du paramètre de population recherché, c'est-à-dire que la formule de l'estimateur du paramètre de population est appliquée à l'échantillon tout entier. On a seulement besoin de la variable de pondération de l'échantillon $WTVAR$ pour ce calcul.

31. Deuxièmement, pour estimer la variance de cet estimateur, on forme plusieurs sous-échantillons ou « répliqués » différents à partir de l'échantillon complet, de telle manière que chaque répliqué reflète le plan d'échantillonnage ainsi que les procédures de pondération et les ajustements apportés à l'échantillon complet. Chaque répliqué se définit par la valeur d'une variable de pondération. Par exemple, $REPWT_j$ est la variable de pondération du répliqué # j , où $j = 1, 2, 3, \dots, G$ (nombre total de répliqués). Une observation de l'échantillon complet a une valeur nulle pour $REPWT_j$ si elle n'est pas incluse dans le répliqué # j et a une valeur positive dans le cas contraire. La somme des valeurs de $REPWT_j$ à travers les observations de l'échantillon complet est une estimation du nombre d'éléments de la population.

32. Troisièmement, on applique la formule de l'estimateur à chaque répliqué pour obtenir une estimation ponctuelle du paramètre de population recherché (l'estimation du répliqué), ce qui donne G estimations du répliqué du même paramètre de population.

33. Quatrièmement, sur la base de la variabilité des G estimations de répliqué, on calcule une estimation de la variance de l'estimateur de l'échantillon complet.

34. Les répliqués peuvent être formés de différentes façons, selon diverses techniques de réplification. Deux des principales approches de la formation de répliqués, chacune avec des variantes, sont la réplification répétée équilibrée [*balanced repeated replication* (BRR)] et la réplification « jackknife » (examinées toutes deux ci-après). La publication d'ensembles de données destinés à faire l'objet d'une estimation de la variance par une méthode de réplification particulière inclut généralement les variables de pondération des répliqués dans ces ensembles de données. Dans ce cas, l'analyste des données secondaires doit utiliser un logiciel d'estimation de la variance comprenant la technique de réplification pour laquelle les pondérations des répliqués de l'ensemble de données ont été établies.

35. Toutefois, on peut vouloir utiliser une technique de réplification pour estimer la variance lorsque les pondérations des répliqués ne figurent pas déjà dans l'ensemble de données. Les logiciels qui comprennent des méthodes d'estimation de la variance calculent également les pondérations des répliqués. Le minimum de variables de conception de l'enquête nécessaire à un logiciel pour fixer les pondérations des répliqués sont : la variable de pondération, $WTVAR$, la variable de stratification, $STRATVAR$, et la variable des UPE, $PSUVAR$ d'une strate donnée. Si l'échantillon complet a été ajusté

pour les cas de non-réponse, ces informations peuvent alors être également acceptées comme entrée par le logiciel pour le calcul des pondérations des réplicats (par exemple, WesVar). On peut toujours calculer soi-même les pondérations des réplicats, mais cela n'est recommandé que pour ceux qui sont rompus aux techniques de réplication.

3. Réplication répétée équilibrée [*balanced repeated replication (BRR)*]

36. La réplication répétée équilibrée (BRR) est une technique particulière de réplication qui peut être utilisée pour des conceptions très générales, telles que l'échantillonnage stratifié à phases multiples. Cependant, elle a été conçue pour le cas particulier de sélection de deux UPE par strate, sélectionnées généralement selon une probabilité inégale avec ou sans remplacement. Elle est également appliquée avec l'approximation *WR* à un plan d'échantillonnage complexe (l'approche UCVE).

37. Avec la réplication BRR, chaque réplicat comprend exactement la moitié des UPE de l'échantillon, à raison d'une UPE par strate; souvent, chaque réplicat est appelé « demi-échantillon ». Le nombre total de cas possibles de réplicats différents est 2^L , où L est le nombre de strates. Cela dit, il n'est pas nécessaire d'utiliser tous les réplicats 2^L , car cela pourrait nécessiter un temps de calcul excessif. En revanche, un ensemble plus petit et « équilibré » de réplicats peut donner la même estimation de la variance que celle que donneraient tous les réplicats. Des réplicats « équilibrés » G sont formés à l'aide d'une matrice Hadamard (Wolter, 1985), de sorte que chaque UPE de l'échantillon et chaque paire d'UPE de deux strates différentes apparaissent dans le même nombre de réplicats. Le nombre minimum G de réplicats nécessaires est le plus petit nombre entier égal ou supérieur à L mais divisible par 4. Par exemple, 49 strates ayant chacune deux UPE sélectionnées nécessiteraient 52 réplicats BRR. Les observations sur les UPE sélectionnées qui ne figurent pas dans le réplicat j ont une valeur nulle pour la variable de pondération du réplicat $REPWT_j$, et les observations sur les UPE sélectionnées qui figurent dans le réplicat j ont une valeur égale au double de leur pondération d'échantillonnage dans l'échantillon complet, quoique cette valeur puisse avoir à être ajustée par les cas de non-réponse et/ou la post-stratification.

38. Fay (Judkins, 1990) a conçu une variante de la technique BRR définie ci-dessus car la réplication BRR standard peut être problématique si l'estimation est souhaitée pour un petit domaine ou pour un ratio de population lorsque le dénominateur compte peu de cas dans l'échantillon complet. Dans la méthode de Fay, les observations concernant les UPE de l'échantillon qui ne sont pas choisies pour le réplicat j ne sont pas annulées comme elles le sont avec la méthode BRR standard. En fait, leur pondération d'échantillonnage est diminuée d'un facteur de multiplication K ($0 \leq K < 1$), tandis que les observations sur les UPE choisies pour le réplicat ont leur pondération renforcée par le facteur de multiplication $(2 - K)$. Le choix de $K = 0$ correspond à la technique BRR standard. Pour la méthode de Fay, on recommande souvent la valeur $K = 0,3$.

4. Techniques de réplication « jackknife » (JK)

39. Les techniques « jackknife » consistent à supprimer une UPE de l'échantillon à la fois pour former des réplicats puis à repondérer chaque réplicat en tant que de besoin de manière à représenter la population de l'échantillon complet. Une UPE peut ne comprendre qu'un seul élément, comme dans le cas de l'échantillonnage aléatoire simple ou de l'échantillonnage aléatoire stratifié, ou plusieurs éléments comme dans le plan d'échantillonnage à approximation *WR*.

40. Considérons d'abord le cas où aucune stratification n'est utilisée avant la sélection des UPE et où chaque UPE de l'échantillon G (comptant approximativement le même nombre d'éléments) ressemble à l'échantillon complet. On forme un total de G réplicats en supprimant une UPE à la fois. Pour le réplicat j dont la variable de pondération est $REPWT_j$, les observations sur l'UPE $\#j$ supprimée ont une valeur nulle pour $REPWT_j$. Chaque observation des UPE restantes (non supprimées) de l'échantillon ont une valeur pour $REPWT_j$ égale à la pondération d'échantillonnage pour cette observation multipliée par le facteur $[G/(G - 1)]$.

41. Considérons un deuxième exemple comprenant L strates dont exactement deux UPE sont sélectionnées pour chaque strate, ce qui correspond au cas examiné ci-dessus pour BRR. La suppression d'une UPE à la fois produirait des réplicats $2L$. Pour chacun de ces réplicats $2L$, l'UPE restant dans chaque strate de l'échantillon où une UPE a été supprimée aurait une pondération d'échantillonnage pour chaque observation multipliée par 2 (et l'UPE supprimée aurait une pondération multipliée par zéro). Toutefois, cette technique n'est généralement utilisée qu'avec L réplicats plutôt que $2L$ réplicats, où une seule UPE de l'échantillon, choisie au hasard, est supprimée de chacune des strates L . Pour les estimateurs linéaires, l'estimateur de la variance correspondant aux L réplicats équivaut algébriquement à l'estimateur de la variance correspondant aux $2L$ réplicats.

42. Le plan d'échantillonnage le plus courant est l'échantillonnage stratifié à phases multiples à L strates (avant sélection des UPE) et à deux UPE ou plus sélectionnées pour chaque strate. Chaque UPE de l'échantillon est supprimée pour former un réplicat; le nombre de réplicats G est égal au nombre total d'UPE de l'échantillon complet (n). Dans la strate h , la valeur de la variable de pondération du réplicat $REPWT_j$ pour chaque observation dans l'UPE supprimée est la variable de pondération de l'échantillon $WTVAR$ multipliée par zéro. La valeur de la variable $REPWT_j$ pour chaque observation restante sur la strate h d'où l'UPE de l'échantillon a été supprimée est la variable de pondération de l'échantillon $WTVAR$ multipliée par le facteur $[n_h/(n_h - 1)]$, où n_h est le nombre d'UPE de la strate h de l'échantillon complet.

5. Quelques erreurs courantes commises par les utilisateurs de logiciels d'estimation de variance

43. Plusieurs logiciels exigent de l'utilisateur qu'il trie les données d'entrée en fonction de certaines des variables de conception de l'enquête : par exemple, *STRATVAR* et *PSUVAR* à l'intérieur de *STRATVAR* (comme il est expliqué au par. 35). L'omission d'un tel tri risque de produire des estimations inexactes de la variance, quoique la plupart des logiciels émettent un message d'erreur si les données ne sont pas triées correctement.

44. Les utilisateurs d'ensembles de données à publier peuvent spécifier des variables de conception d'enquête incorrectes par suite d'un examen inadéquat de la documentation de l'enquête par sondage. Une pondération d'échantillon incorrectement spécifiée se traduit par des estimateurs biaisés et des variances estimatives incorrectes, c'est-à-dire que toutes les analyses ainsi produites sont erronées. Si la variable de pondération de l'échantillon est correcte mais pas la variable de pondération et/ou d'UPE, les estimations ponctuelles seront correctes, mais pas les estimations de la variance.

45. Certains ensembles de données publiés comportent de multiples fichiers de données assortis de différentes variables de conception de l'enquête pour différents fichiers. Ces différents fichiers de données peuvent avoir différentes unités à analyser, par exemple la personne, le ménage ou la famille, de sorte qu'il faut prendre grand

soin lors de l'interprétation des résultats. Certaines variables d'enquête ne peuvent se mesurer que sur un sous-échantillon probabiliste de l'échantillon complet, et nécessitent une variable de pondération de l'échantillon différente des variables mesurées sur l'échantillon complet. Il importe de lire attentivement la documentation de toutes les enquêtes par sondage, que le plan d'échantillonnage soit simple ou particulièrement complexe.

D. Comparaison des logiciels d'estimation de variance

46. On peut trouver des liens Internet sur tout un ensemble de logiciels d'enquêtes par sondage, y compris sur les huit logiciels examinés dans cet article, sur le site Web d'information www.fas.harvard.edu/~stats/survey-soft/survey-soft.html. [Voir également Carlson (1998) pour un examen des logiciels pour données d'enquêtes par sondage complexes.] On notera que SPSS ne figure pas parmi les logiciels examinés. Au début de 2003, SPSS ne pouvait pas servir à l'estimation de la variance d'enquêtes par sondage complexes, mais un module supplémentaire est sorti à la fin de 2003, alors que ce chapitre était à l'impression.

Tableau XXI.1

Comparaison de PROCS selon cinq logiciels : pourcentage estimatif et nombre de femmes qui sont séropositives, et estimation de l'erreur type, femmes ayant accouché récemment, Burundi, 1988-1989

Logiciel et PROC	% séropos	s.e. du % séropos	95% CI % séropos	Nombre séropos	s.e. # séropos	95% CI # séropos
SAS 8.2 MEANS ^a Sans pondération	74,88% wrong	2,12% Wrong	N-APP	N-APP	N-APP	N-APP
SAS 8.2 MEANS ^b Avec pondération	67,20%	2,30% Wrong	N-APP	N-APP	N-APP	N-APP
SAS 8.2 SURVEYMEANS	67,20%	3,83%	59,38%, 75,02%	142 485	8 848,10	124 415, 160 556
SUDAAN 8.0 CROSTAB et DESCRIPT Taylor et BRR	67,20%	3,83%	N-AV	142 485	8 848,10	N-AV
STATA 7.0 Svymean	67,20%	3,83%	58,38%, 75,02%	N-AV	N-AV	N-AV
STATA 7.0 Svytotal	N-AV	N-AV	N-AV	142 485	8 848,10	124 415, 160 556
Epi-Info 6.04d CSAMPLE ^c	67,20%	3,83%	59,70%, 74,71% ^{c/}	N-AV	N-AV	N-AV
WesVar 4.2	67,20%	3,83%	59,38%, 75,02%	142 485	8 848,10	124 415, 160 556

Note : Abréviations utilisées : CI = Intervalle de confiance, N-APP = sans objet, N-AV = non connu, s.e. = erreur type.

^a Analyse incorrectement spécifiée; ignore la pondération d'échantillonnage, la mise en grappe et la stratification.

^b Analyse incorrectement spécifiée, inclut la pondération d'échantillonnage, mais pas la mise en grappe et la stratification

^c L'Intervalle de confiance donné par Epi-Info 6.04d est plus étroit que celui donné par les autres logiciels. Epi-Info 6.04d a utilisé $z = t = 1,96$ pour construire l'intervalle de confiance de 95 %, alors que les autres logiciels ont utilisé $t = 2,042$ tiré de la distribution Student-t avec un ddf (degré de liberté en dénominateur pour l'enquête par sondage, calculé comme le nombre d'UPE moins le nombre de pseudo-strates). Il est préférable d'utiliser le ddf effectif de l'enquête.

47. Le reste de ce chapitre est consacré à l'examen et à la comparaison de huit logiciels d'estimation de variance de données d'enquêtes par sondage complexes : SAS, SUDAAN, STATA, Epi-Info, WesVar, PC-CARP, CENVAR et IVEware. Les cinq premiers sont illustrés à l'aide d'analyses descriptives utilisant les données d'une enquête par sondage effectuée en 1989 au Burundi; les proportions, moyennes et totaux de population sont estimés et les domaines sont comparés pour ces paramètres. Les résultats des analyses de l'enquête effectuée au Burundi sont résumés dans ce chapitre, au tableau XXI.1, et des tableaux détaillés et des exemples annotés de programmes et de résultats concernant chaque logiciel sont donnés dans l'annexe figurant sur le CD-ROM. Ces exemples annotés peuvent aider les utilisateurs à apprendre comment utiliser les cinq premiers logiciels d'estimation de variance.

48. Parmi les cinq logiciels illustrés à l'aide des données de l'enquête effectuée au Burundi, trois (STATA, SAS et Epi-Info) présentent des procédures d'enquête par sondage à l'intérieur d'un logiciel de statistiques générales. Tous trois utilisent la linéarisation par série de Taylor pour l'estimation de la variance. Les deux autres (WesVar et SUDAAN) ont été conçus spécialement pour l'estimation de la variance dans les enquêtes par sondage. WesVar utilise les méthodes de réplication et SUDAAN offre à la fois les méthodes de linéarisation par série de Taylor et de réplication.

49. Trois autres logiciels (PC-CARP, CENVAR et IVEware) sont examinés, mais ne sont pas illustrés à l'aide des données de l'enquête du Burundi. PC-CARP et CENVAR utilisent tous deux la linéarisation par série de Taylor pour l'estimation de la variance. IVEware utilise à la fois les méthodes de linéarisation par série de Taylor et de réplication.

50. Les huit logiciels examinés ici comprennent un grand nombre, mais pas la totalité, des options offertes pour l'estimation de la variance dans les enquêtes par sondage. Trois (Epi-Info, CENVAR et WesVar 2) ont été choisis parce qu'ils offrent des analyses descriptives élémentaires et peuvent être téléchargés gratuitement du Web, ce qui est intéressant pour les analystes dont le budget d'achat de logiciels est limité, voire inexistant. Deux (PC-CARP et WesVar 4) ont été choisis parce que, sans être gratuits, ils sont peu coûteux par comparaison avec d'autres options et offre des analyses descriptives ainsi que des régressions linéaire et logistique fondées sur la conception de l'enquête. Deux logiciels d'un prix modique (SUDAAN et STATA) ont été choisis parce qu'ils offrent, en plus d'analyses descriptives, des choix détaillés de modèles de régression fondés sur la conception. Bien que coûteux, SAS a été choisi en raison de sa prépondérance en ce qui concerne la gestion et l'analyse des données et son PROCs relativement nouveau pour l'analyse des données d'enquêtes par sondage. Enfin, le logiciel IVEware (version bêta) sorti récemment a été choisi parce qu'il offre des analyses descriptives détaillées et des modèles de régression fondés sur la conception, ainsi que des procédures d'imputation multiple. IVEware est gratuit (et peut être téléchargé du Web) mais fonctionne comme application du SAS (et nécessite donc ce logiciel).

51. Le tableau XXI.2 résume diverses caractéristiques des huit logiciels, y compris les plans d'échantillonnage couverts, les méthodes d'estimation de la variance et les types d'analyse.

Tableau XXI.2
Attributs de huit logiciels permettant l'estimation de la variance pour les données
d'enquêtes par sondage complexes

ATTRIBUT	SAS 8.2	SUDAAN 8.0	STATA 8.0	Epi-Info 6.04d	WesVar 4.2	PC-CARP	CENVAR	IVeware
Série de Taylor	Oui	Oui	Oui	Oui	Non	Oui	Oui	Oui Desc
Méthodes de réplification BRR et JK	Non	BRR JK	Non	Non	BRR JK	Non	Non	Modèles JK
Pondération des réplcats formés	Non	Non-BRR Oui-JK	No	Non	Oui BRR/JK	Non	Non	Oui JK
Ensemble de données d'entrée	SAS	SAS, SPSS, ASCII	STATA	Epi-Info	SAS, SPSS, STATA, ASCII, ODBC	ASCII	ASCII	SAS
Estimation du total	Oui	Oui	Oui	Non	Oui	Oui	Oui	Non
CI sur le total	Oui	Non	Oui	Non	Oui	Oui	Oui	Non
LC sur les totaux	Non	Oui	Oui	Non	Oui	Oui	Oui	Non
Estimation de la moyenne	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui
CI sur la moyenne	Oui	Non	Oui	Oui-étroit	Oui	Oui	Oui	Oui
LC sur la moyenne	Non	Non	Oui	Oui	Oui	Oui	Oui	Oui
Estimation des proportions	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui
CI sur les proportions	Oui	Non	Oui	Oui-étroit	Oui	Oui	Oui	Oui
LC sur les proportions	Non	Oui	Oui	Oui- erreur	Oui	Oui	Oui	Oui
Estimation des ratios	Oui	Oui	Oui	Non	Oui	Oui	Oui	Non
CI sur les ratios	Oui	Non	Oui	Non	Oui	Oui	Oui	Non
LC sur les ratios	Non	Oui	Oui	Non	Oui	Oui	Oui	Non
Analyse des domaines	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui
Comparaison des domaines	Non-8.2 Oui-9.0	Oui	Oui	Oui	Oui	Oui	Oui	Oui
Analyse des sous-ensembles de population	Non	Oui	Oui	Non	Oui	Oui	Oui	Oui
Taux/moyennes normalisés	Non-8.2 Oui-9.0	Oui	Oui	Non	Oui	Non	Non	Non
Tests de chi-carré	Non-8.2 Oui-9.0	Oui	Oui	Non	Oui	Oui	Non	Non
Régression logistique	Non	Oui	Oui	Non	Oui	Oui	Non	Oui
Ratio d'anomalies	Non	Oui	Oui	Oui	Oui	Oui	Non	Oui
Ratio de risque	Non	Oui s	Oui	Oui	Oui	Non	Non	Non
Régression linéaire	Oui	Oui	Oui	Non	Oui	Oui	Non	Oui
Modèles de régression additionnels	Non	Oui	Oui	Non	Non	Non	Non	Oui
Décrit de nombreuses phases d'échantillonnage	Non	Oui	Non	Non	Non	Non	Non	Non
Effet de conception	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui
Logiciel d'essai gratuit	Non	Non	Non	NA Gratuit	Oui	Non	NA Gratuit	NA Gratuit
Logiciel général de statistiques	Oui	Non	Oui	Oui	Non	Non	Non	Non

ATTRIBUT	SAS 8.2	SUDAAN 8.0	STATA 8.0	Epi-Info 6.04d	WesVar 4.2	PC-CARP	CENVAR	IVEware
Capacité de gestion des données	Oui	Non	Oui	Oui	Oui	Non	Non	Non
Fonctionne avec programmes d'entrée	Oui	Oui	Oui	Non-6.04d Oui-2002	Non	Non	Non	Oui
Fonctionne sur commandes brèves	Non	Non	Oui	Non	Non	Non	Non	Non
Fonctionne par sélection sur menu	Non	Non	Non	Oui	Oui	Oui	Oui	Non
Trie les données par strate et par UPE	Non	Oui	Non	Oui	Non	Oui	Oui	Non
Formation offerte par l'inventeur	Oui	Oui	Oui	Non	Oui	Non	Non	Non
Manuel écrit/en ligne	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui
Cours sur les procédures d'enquête	Non	Non	Non	Non	Oui	Non	Oui	Non
Coût	Élevé	Moyen	Moyen	Gratuit	Faible V4 Gratuit V2	Faible	Gratuit	Gratuit
Redevance annuelle de renouvellement	Élevé	Moyen	Aucune	Aucune	Aucune	Aucune	Aucune	Aucune
Entre les données	Non	Non	Non	Non	Non	Oui	Non	Oui

Abréviations utilisées : ASCII = American Standard Code for Information Interchange, BRR = *balanced repeated replication*, CI = intervalle de confiance, JK = « jackknife », LC = contraste linéaire, NA = non connu, ODBC = *Open DataBase Connectivity*, V = version.

E. L'ensemble de données de l'enquête par sondage effectuée au Burundi

52. Tous les exemples numériques présentés dans ce chapitre sont fondés sur un ensemble de données tiré d'une enquête par sondage sur la couverture d'une campagne de vaccination au tétanos toxoïde lancée en 1989 au Burundi. On trouvera ci-après un bref aperçu de la conception de cette enquête, dont une description plus détaillée figure à la section I de l'annexe sur CD-ROM. Pour tous renseignements complémentaires sur la méthodologie de cette enquête et ses résultats publiés, voir le rapport sur le Programme élargi de vaccination (PEV) [1996] de l'Organisation mondiale de la santé (OMS).

1. Population étudiée et paramètres de population

53. La population considérée pour cette enquête est celle des femmes du Burundi qui ont accouché entre Pâques 1988 et février-mars 1989. Le paramètre de population étudié est le pourcentage (ou la proportion) de femmes qui ont testé positif à l'antitoxine du tétanos, assurant ainsi la protection du nouveau-né contre le tétanos néonatal.

2. Plan d'échantillonnage et collecte de données

54. Le plan d'échantillonnage était une version modifiée suggérée par Brogan *et al.* (1994) de la méthodologie d'enquête par sondage en grappe conçue par l'OMS pour son Programme élargi de vaccination. Cette version modifiée produit un échantillon de logements ou d'unités d'habitation et, par conséquent, un échantillon probabiliste de femmes que ne produirait pas nécessairement la méthodologie standard d'échantillonnage en grappe du PEV de l'OMS (ibid).

55. Le Burundi a été stratifié en deux zones géographiques, la capitale Bujumbura (strate urbaine) et le reste du pays (strate rurale). Les unités primaires d'échantillonnage (UPE) étaient des zones géographiques, les *collines* dans la strate rurale et les *quartiers* ou *avenues* dans la strate urbaine. Le cadre d'échantillonnage des UPE de chaque strate était fondé sur la proximité géographique. Un échantillonnage systématique fondé sur une probabilité proportionnelle à la taille estimative [*probability proportional to estimated size* (ppes)] a été utilisé pour sélectionner 30 UPE par strate. Étant donné que 96 % de la population étudiée vivent en milieu rural et que le même nombre d'UPE a été alloué à chaque strate, les femmes de la zone urbaine se sont trouvées sensiblement sur-représentées. La disposition spécifique des UPE sur le cadre d'échantillonnage, combinée à un échantillonnage pps systématique des UPE, donne une stratification géographique implicite à l'intérieur de chaque strate.

56. D'autres phases d'échantillonnage probabiliste ont été organisées à l'intérieur des UPE pour produire un échantillon de logements occupés. Toutes les femmes admissibles pour l'enquête vivant dans un logement sélectionné ont été retenues pour l'échantillon. Le fait de tester positif à l'antitoxine du tétanos a été déterminée par le prélèvement d'une goutte de sang au bout du doigt. Le taux de réponse à l'enquête a été essentiellement de 100 %, ce qui est inhabituel. Au total, 206 femmes de la strate urbaine et 212 femmes de la strate rurale ont été interrogées.

3. Procédures de pondération et cadre d'estimation de la variance

57. La variable de pondération de l'échantillon W de l'ensemble de données du Burundi a été remplacée par la variable $W2$, de sorte que la valeur de $W2$ pour un répondant R est une estimation du nombre de femmes de la population étudiée, représentée par cette lettre R . La valeur de $W2$ est une approximation utilisée uniquement pour illustrer l'estimation des totaux de population obtenue avec les divers logiciels. Il ne faudrait pas tirer de conclusions de fond des analyses de ce chapitre en ce qui concerne les totaux de population à partir des femmes aptes à figurer dans l'enquête de 1989 au Burundi. Il convient de noter que les proportions et les moyennes estimatives indiquées dans ce chapitre concordent avec les résultats publiés précédemment avec cet ensemble de données (Programme élargi de vaccination, 1996) car la variable révisée $W2$ est un multiple scalaire de W utilisé pour des analyses précédentes. La valeur de $W2$ était de 959,3 pour les femmes de l'échantillon rural et de 42,0 pour celles de l'échantillon urbain, ce qui s'explique par la forte sur-représentation des femmes urbaines. Une version approximative du plan d'échantillonnage du Burundi est donnée par la description courante de WR pour l'estimation de la variance, c'est-à-dire l'approche UCVE avec de faibles fractions d'échantillonnage de première phase.

58. Comme les UPE ont été implicitement stratifiées, le plan d'échantillonnage de chacune des strates urbaine et rurale a été considéré comme deux UPE sélectionnées à partir de chacune des 15 pseudo-strates. Il est jugé préférable de décrire le plan d'échantillonnage comme un total de 30 pseudo-strates plutôt que comme un ensemble de 2 strates, chacune avec 30 UPE sélectionnées, parce qu'on obtient ainsi une estimation moins biaisée de la variance, qui tient compte de la stratification implicite. La variable de pseudo-stratification $PSTRA$ a été codée de 1 à 30 et la variable de pseudo-UPE $PPSU$ a été codée 1 ou 2 dans chaque strate.

59. Lorsque l'on utilise la linéarisation par série de Taylor pour estimer la variance, on n'a besoin que des variables $W2$, $PSTRA$ et $PPSU$. En revanche, quand on utilise les techniques de réplification, il faut les pondérations des répliqués. On a utilisé $WesVar$ pour calculer les pondérations des répliqués de BRR à partir des variables $W2$,

PSTRA et *PPSU*. On a utilisé ces pondérations à la fois avec *WesVar* et *SUDAAN* pour estimer la variance en utilisant *BRR*.

4. Trois exemples d'analyses de données d'enquêtes

60. L'annexe contient des analyses annotées des données pour les trois exemples ci-dessous, effectuées à l'aide de cinq logiciels pour données d'enquêtes par sondage (sections II-VI). Les exemples ci-dessous illustrent des analyses descriptives et analytiques courantes effectuées sur de telles données, à savoir : *a*) estimation des proportions, des totaux et des moyennes pour l'ensemble de la population et pour certains domaines ou strates; et *b*) comparaisons des domaines ou des strates à partir des moyennes ou des proportions. La population étudiée est constituée par les femmes qui étaient aptes à figurer dans l'enquête au Burundi au début de 1989.

- Exemple 1 : Estimer le nombre de femmes (total de population) et le pourcentage des femmes (proportion/pourcentage de population) qui ont testé positif (variable d'*IMMUNITÉ*, 1 = positif; 2 = négatif). La variable *SANG* (1 = positif; 2 = négatif) est un « recode » de la variable d'*IMMUNITÉ*.
- Exemple 2 : Estimer les paramètres de population de l'exemple 1 pour les femmes urbaines et rurales (variable *RUR_URB*, 1 = rurales; 2 = urbaines). Déterminer si la caractéristique rurale/urbaine est statistiquement indépendante de la séropositivité (*IMMUNITÉ*).
- Exemple 3 : Estimer les unités internationales moyennes d'antitoxine par millilitre (ml) [*IUML*], pour la population des femmes étudiées et pour les catégories rurale/urbaine. Déterminer s'il y a un lien entre l'appartenance à la catégorie rurale ou urbaine et l'*IUML* moyenne.

Note : L'estimation des unités internationales moyennes d'antitoxine par millilitre (ml) [*IUML*] peut donner des résultats trompeurs en raison de la distribution inégale de cette variable. Il pourrait être préférable d'utiliser la valeur médiane ou de remplacer l'*IUML* avant l'analyse, par exemple par le logarithme naturel de l'*IUML*. Dans ce chapitre, la moyenne de l'*IUML* (sans remplacement) est censée indiquer les capacités des cinq logiciels, et non pas illustrer les résultats concrets concernant l'*IUML*.

F. Utilisation des procédures d'enquête sans échantillon pour analyser les données d'enquêtes avec échantillonnage

61. La présente section montre que la mauvaise utilisation de formules d'échantillonnage aléatoire simple pour l'analyse de données d'enquêtes complexes peut déboucher sur des estimations ponctuelles biaisées et des estimations (généralement trop faibles) d'erreurs types biaisées [voir Brogan (1998 et sous presse) pour une autre illustration]. N'importe quel logiciel de statistiques aurait pu être utilisé pour cette illustration, et l'on aurait obtenu des réponses comparables à celles obtenues avec *SAS* (utilisé dans cette section).

62. Le paramètre de population à estimer est la proportion de femmes dans la population étudiée qui sont séropositives. La variable indicateur *BLOOD* (sang) est calculée et codée comme 1 = séropositive et 0 = séronégative. Autrement dit, la moyenne de l'indicateur *BLOOD* est la proportion de femmes séropositives. Dans *SAS*, *PROC MEANS* estime la moyenne de *BLOOD* à 0,74880, avec une erreur estimative de 0,02124 (rang 1 du tableau XXI.1). Ces deux calculs sont faussés parce que la mauvaise application de *PROC MEANS* ne tient pas compte de la variable de pon-

dération de l'échantillonnage, de la variable UPE et de la variable de stratification pour le calcul de l'erreur type estimative de l'estimation ponctuelle.

63. Dans SAS, PROC MEANS est donc utilisé avec la variable de pondération de l'échantillonnage W2 sur un état des PONDÉRATIONS. La moyenne de BLOOD est estimée à 0,67203, avec une erreur type estimative (pondérée) de 0,02299 (rang 2 du tableau XXI.1). Dans cette analyse, PROC MEANS donne une estimation ponctuelle appropriée de la proportion de population. Cependant, la mauvaise application de PROC MEANS donne une erreur type estimative faussée car elle ne tient pas compte de la variable UPE et de la variable de stratification.

64. Enfin, PROC SURVEYMEANS dans SAS est utilisé pour donner une analyse appropriée des données d'une enquête par sondage complexe (on trouvera les indications sur la façon d'utiliser SURVEYMEANS dans la section suivante). L'estimation ponctuelle de la proportion de population est de 0,67203, avec une erreur type estimative de 0,03830 (rang 3 du tableau XXI.1). PROC SURVEYMEANS tient compte de la variable de pondération de l'échantillonnage pour calculer l'estimation ponctuelle et la pondération de l'échantillonnage, la variables UPE et la variable de stratification pour calculer l'erreur type estimative.

65. Une comparaison de ces trois analyses au tableau XXI.1 montre que l'estimation ponctuelle (inexacte) non pondérée de 0,7488 (74,88 %) diffère très sensiblement de l'estimation ponctuelle (exacte) pondérée de 0,6720 (67,20 %). L'estimation ponctuelle non pondérée est trop élevée parce qu'une plus forte proportion de femmes urbaines que de femmes rurales sont séropositives (comme on le verra plus loin dans ce chapitre), et que les femmes urbaines sont sur-représentées dans l'échantillon car elles ont été sur-sélectionnées : elles constituent environ la moitié de l'échantillon mais ne représentent que 4 % de la population étudiée. Ainsi, dans une analyse non pondérée du pays, les femmes urbaines ont beaucoup plus d'influence qu'elles ne devraient avoir et faussent à la hausse l'estimation de la proportion de population.

66. Une comparaison des deux analyses qui donnent l'estimation ponctuelle pondérée correcte montre que, même avec une indication WEIGHT (PONDÉRATION), la mauvaise application de PROC MEANS dans SAS sous-estime sérieusement l'erreur type et donne un calcul inexact de 0,02299 (2,30 %) par rapport au calcul exact de PROC SURVEYMEANS de 0,03830 (3,83 %). Cela est dû principalement au fait que PROC MEANS, avec ou sans indication WEIGHT, ne tient pas compte de la mise en grappe des femmes des UPE de l'échantillon, tandis que SURVEYMEANS en tient compte pour l'estimation de la variance. Comme le coefficient de corrélation intra-classe est positif pour la plupart des variables mesurées dans les enquêtes par sondage complexes, les bonnes procédures d'estimation de la variance qui tiennent compte de la mise en grappe produisent de plus grosses erreurs types estimatives.

67. En général, on obtient des estimations ponctuelles faussées des paramètres de population si les données de l'enquête ne sont pas analysées avec la bonne variable de pondération de l'échantillon. En outre, même si la variable de pondération de l'échantillon est incorporée dans l'analyse et produit des estimations ponctuelles appropriées des paramètres de population, les erreurs types sont généralement sous-estimées lorsque les éléments de l'échantillon sont mis en grappe dans les données de l'enquête et que l'estimation de la variance ne tient pas compte de cette mise en grappe. La sous-estimation des erreurs types engendre des intervalles de confiance qui sont trop étroits et des tests statistiques de signification dont les valeurs-p sont trop faibles; autrement dit, le niveau de signification statistique est surestimé.

68. L'ampleur de la sous-estimation de la variance causée par le fait de ne pas tenir compte de la mise en grappe des données de l'enquête par sondage est exprimée

de façon approximative par l'expression $[1 + \rho (b - 1)]$ où ρ est le coefficient de corrélation intra-classe entre les éléments de la population et b est le nombre moyen d'éléments de l'échantillon par grappe (UPE) [voir chapitre VI]. Par exemple, si la valeur de l'expression est 2, le fait de tenir compte de la mise en grappe double approximativement la variance estimative que l'on obtiendrait en ne tenant pas compte de la mise en grappe. Il convient de noter que la variable UPE au Burundi désignée par les lettres *PPSU* indique pour le logiciel quels éléments de l'échantillon sont mis en grappe à l'intérieur de l'UPE de l'échantillon (pour une strate donnée).

69. Outre l'impact de la mise en grappe sur la variance estimative, une variation substantielle de la valeur de la variable de pondération de l'échantillonnage à travers les répondants augmente la variance estimative. Autrement dit, si l'on ne tient pas compte dans l'analyse de la variable de pondération de l'échantillonnage, on sous-estime l'erreur type estimative (et l'estimateur du paramètre de population est faussé).

G. Les procédures d'enquête par sondage avec SAS 8.2

1. Aperçu général de SURVEYMEANS et de SURVEYREG

70. La version 8.2 de SAS contient deux procédures récemment mises au point (elles sont apparues pour la première fois dans V 8.0) pour l'analyse des données d'enquêtes par sondage : SURVEYMEANS et SURVEYREG. SAS donne la description courante du plan d'échantillonnage *WR* qui nécessite les trois variables essentielles à la conception de l'enquête. Les termes de correction de la population finie peuvent être appliqués aux conceptions d'échantillonnage à phase unique tels l'échantillonnage aléatoire stratifié et l'échantillonnage aléatoire simple. On utilise la linéarisation par série de Taylor pour estimer la variance. SAS V9 contient deux nouvelles PROCs pour les données d'enquêtes par sondage complexes : SURVEYFREQ pour l'analyse des variables catégoriques et SURVEYLOGISTIC pour la régression logistique. D'autres procédures SAS sont à l'étude pour les données d'enquêtes par sondage.

71. La syntaxe pour spécifier les variables de conception d'enquête pertinentes *WR* est la même pour SURVEYMEANS que pour SURVEYREG. On utilise le mot clé STRATA pour désigner la variable de stratification, le mot clé CLUSTER pour désigner la variable des UPE et le mot clé WEIGHT pour désigner la variable de pondération de l'échantillonnage (comme dans d'autres procédures SAS telles que MEANS). Ces appellations, qui conviennent pour une enquête donnée, doivent figurer dans chaque procédure d'enquête par sondage SAS et ne changent généralement pas tant que se poursuit l'analyse de l'ensemble de données de la même enquête par sondage. Pour l'ensemble de données du Burundi, les appellations ci-dessous décrivent la conception de l'enquête par sondage pour SAS PROC SURVEYMEANS ou PROC SURVEYREG :

```
STRATA PSTR  
CLUSTER PPSU  
WEIGHT W2
```

72. Si l'instruction STRATE (STRATA) manque, SAS suppose que le plan d'échantillonnage ne comportait pas de stratification des UPE avant l'échantillonnage de première phase. Si l'instruction GRAPPE (CLUSTER) manque, SAS suppose que les éléments de l'échantillon ne sont pas mis en grappe, c'est-à-dire que chaque grappe de l'échantillon contient exactement un élément, ce qui veut dire que les éléments ont été sélectionnés à la première (et unique) phase de l'échantillonnage, et que l'on a utilisé l'échantillonnage aléatoire simple ou l'échantillonnage aléatoire stratifié. Si

l'instruction PONDÉRATION (WEIGHT) manque, SAS suppose que chaque R a la même valeur pour la variable de pondération, et SAS affecte la valeur 1,0 à la variable de pondération. Si les trois instructions (STRATA, CLUSTER, WEIGHT) manquent, cela revient à spécifier un échantillonnage aléatoire simple à partir d'une population infinie, l'hypothèse utilisée pour la plupart des PROCs de SAS non utilisées pour les enquêtes.

2. SURVEYMEANS

73. Cette procédure sert à estimer les moyennes et les totaux de population pour des variables continues et les proportions et les totaux de population pour les variables catégoriques à l'aide des données d'enquêtes par sondage. Les erreurs types estimatives et les coefficients de variation sont fournis pour toutes les estimations ponctuelles, ainsi que pour les intervalles de confiance pour les paramètres de population. Des statistiques spécifiques peuvent être demandées sur l'instruction PROC, ou l'on peut prendre l'impression par défaut pour les statistiques, ou l'on peut utiliser ALL sur l'instruction PROC pour obtenir toutes les statistiques qui peuvent être calculées à l'aide de SURVEYMEANS.

74. Les variables à analyser (continues et catégoriques) figurent sur l'instruction VAR. L'instruction CLASS donne la liste des variables de l'instruction VAR qui sont catégoriques; SAS suppose alors que toutes les autres variables figurant sur l'instruction VAR sont continues.

75. On utilise l'instruction DOMAIN avec une ou plusieurs variables catégoriques pour spécifier les domaines d'analyse de toutes les variables de l'instruction VAR. SAS fournit automatiquement des analyses pour la population marginale, autrement dit, la population tout entière en plus des analyses par domaine. Un programme sans instruction DOMAIN ne donne d'estimations que pour la population entière. Bien que l'on puisse utiliser l'instruction BY dans SURVEYMEANS pour obtenir des estimations pour les domaines, cela n'est pas recommandé pour les données d'une enquête par sondage parce que l'on n'utilise pas les formules appropriées pour l'estimation de la variance lorsque l'on utilise l'instruction BY. Utiliser l'instruction DOMAIN pour l'analyse des domaines.

76. SAS V8.2 n'a pas d'instruction qui permette d'analyser une fraction donnée de la population, par exemple seulement les femmes âgées. Toutefois, on peut effectuer des analyses de certaines fractions de la population en définissant d'abord une variable indicatrice, par exemple *OLDERFEM*, qui indique si l'élément de l'échantillon appartient à la fraction en question. On peut ensuite utiliser l'instruction DOMAIN *OLDERFEM* pour obtenir les analyses désirées, et ignorer les résultats de SAS pour les éléments de l'échantillon qui ne sont pas des femmes âgées. Ne pas utiliser l'instruction SAS IF pour limiter l'ensemble de données aux femmes âgées avant de recourir à PROC SURVEYMEANS, car les erreurs types peuvent être calculées de façon incorrecte dans la mesure où SURVEYMEANS ne connaît pas nécessairement le nombre complet de strates et d'UPE retenues pour l'enquête par sondage.

3. SURVEYREG

77. Cette procédure effectue une régression linéaire pour les données d'une enquête par sondage selon l'approche fondée sur la conception (Korn et Graubard, 1999); autrement dit, l'analyse tient compte des variables de conception de l'enquête. De même que pour la régression linéaire pour données autres que celles d'une enquête, la variable dépendante est continue (ou considérée comme telle) et les variables

indépendantes peuvent être un mélange de variables continues et de variables catégoriques. L'instruction MODEL comprend la variable dépendante et toutes les variables indépendantes. Toute variable catégorique de l'instruction MODEL doit aussi figurer dans l'instruction CLASS, et celle-ci doit précéder l'instruction MODEL dans le programme SAS. SURVEYREG forme des variables indicatrices fictives (codées 1 ou 0) pour les variables indépendantes catégoriques, la valeur codée la plus élevée de la variable étant définie comme le groupe de référence. Les autres options de SURVEYREG, ainsi que leurs résultats, sont semblables à la régression linéaire (hors enquête) de SAS.

78. La version 8.2 de SAS n'a pas de procédures d'enquête par sondage pour comparer les domaines pour les moyennes ou les proportions, mais ces capacités sont à l'étude. On peut citer comme exemple de question pour ce type de situation la question suivante : les femmes rurales et urbaines de la population étudiée au Burundi diffèrent-elles en unités *IUML* moyennes ou en proportion de femmes séropositives ? SURVEYFREQ de V9.0 peut servir à effectuer un test de chi-carré sur les deux variables résidence (rurale/urbaine) et séropositivité (oui/non). En attendant que les procédures de comparaison de domaines soient pleinement développées dans SAS pour les données d'enquêtes par sondage, SURVEYREG peut être utilisé de la façon suivante pour comparer les domaines.

79. Si l'on désire comparer les femmes rurales et urbaines de la population étudiée du point de vue de l'*IUML* moyenne, utiliser l'instruction MODEL dans SURVEYREG avec la variable continue *IUML* comme variable dépendante et la variable domaine rural/urbain comme variable catégorique indépendante. Une partie des résultats types de SURVEYREG est un test de l'hypothèse nulle selon laquelle le coefficient de régression de la population pour les domaines rural/urbain (avec un degré de liberté) est égal à zéro. Cette hypothèse nulle concernant le coefficient de régression équivaut à l'hypothèse nulle selon laquelle les femmes rurales et urbaines de la population étudiée ont la même *IUML* moyenne.

80. Si l'on désire comparer la proportion des femmes séropositives entre les femmes urbaines et les femmes rurales de la population étudiée (variable binaire), utiliser la variable indicatrice *BLOOD* (1 = séropositif, 0 = séronégatif) comme variable dépendante (on notera que *BLOOD* n'est autre qu'une nouvelle appellation de la variable *IMMUNE* où 1 = séropositif, et 0 = séronégatif). Dans l'instruction MODEL de SURVEYREG, définir *BLOOD* comme la variable dépendante et la variable domaine (rural/urbain) comme la variable catégorique indépendante. L'hypothèse nulle selon laquelle le coefficient de régression est zéro équivaut à l'hypothèse nulle selon laquelle la proportion de femmes séropositives est la même parmi les femmes rurales et urbaines de la population étudiée.

4. Exemples numériques

81. La section II de l'annexe sur le CD-ROM illustre l'application de SURVEYMEANS et SURVEYREG aux trois exemples du paragraphe 60. Un examen des programmes SAS annotés (écrits par les utilisateurs) et des résultats SAS annotés devrait préparer les lecteurs à écrire leur propres programmes SAS pour SURVEYMEANS et SURVEYREG et à interpréter les résultats.

82. Le tableau XXI.1, rang 3, résume les résultats de SURVEYMEANS présentés dans la section II de l'annexe pour l'estimation du pourcentage et du nombre de femmes dans la population étudiée du Burundi qui sont séropositives, avec une erreur type estimative et un intervalle de confiance; la plupart de ces résultats ont été examinés à la section F du présent chapitre. Le tableau XXI.3, rang 1 (annexe, section VII sur

le CD-ROM), résume les résultats de SURVEYMEANS pour l'estimation des pourcentages de femmes séropositives dans chacun des deux domaines (rural et urbain), qui sont respectivement de 66,51 % et 83,50 %. Le tableau XXI.4, rang 1 (annexe, section VII sur le CD-ROM), résume les résultats de SURVEYREG qui compare les femmes rurales et urbaines, produisant une valeur-t de - 3,52, avec une valeur-p de 0,0014 pour vérifier l'hypothèse nulle selon laquelle le pourcentage de femmes séropositives n'est guère différent entre les femmes rurales et urbaines. Ainsi, les femmes rurales et urbaines de la population étudiée diffèrent en pourcentage de femmes séropositives : les femmes urbaines ont un taux de séropositivité plus élevé que les femmes rurales.

5. Avantages/inconvénients/coût

83. Si l'on est déjà utilisateur de SAS/STAT, les procédures d'enquête par sondage de SAS sont disponibles sans dépense supplémentaire et utilisent une syntaxe déjà connue. De plus, on dispose de tous les moyens offerts par SAS pour la gestion des données et la formation de nouvelles variables. L'appui technique et la documentation nécessaires pour les procédures d'enquête par sondage sont accessibles à travers l'appui offert au système SAS. Comparé à celui des autres logiciels d'enquête par sondage examinés, le coût de SAS est élevé.

84. SAS 8-2 n'a pas la capacité de comparer des domaines entre eux, quoique SURVEYREG puisse servir de solution temporaire pour ce type d'analyse. L'addition de SURVEYFREQ dans V9.0 permet des comparaisons de domaines pour les variables catégoriques.

85. SAS utilise seulement la linéarisation par série de Taylor pour l'estimation de la variance. Pour l'échantillonnage en grappe à phases multiples stratifié, il n'accepte que la description du plan d'échantillonnage courant *WR*. Toutefois, il peut incorporer les termes *fpc* dans un échantillonnage aléatoire stratifié à phase unique ou un échantillonnage aléatoire simple.

86. La capacité de SAS 8.2 pour l'analyse des données d'une enquête par sondage est simple et descriptive et peut convenir aux besoins d'analyse de nombreux utilisateurs. L'addition de SURVEYFREQ dans V9.0 permet de disposer d'une capacité descriptive et analytique pour les variables catégoriques. Les procédures d'enquête par sondage qui sont encore au stade de l'élaboration, par exemple la régression logistique, devraient rendre SAS plus comparable à l'avenir aux autres logiciels qui permettent des analyses complètes des données d'enquêtes par sondage.

H. SUDAAN 8.0

1. Aperçu général de SUDAAN

87. SUDAAN (Research Triangle Institute, 2001) est un logiciel spécialisé conçu à l'origine pour l'analyse de données d'enquêtes par sondage complexes mais généralisé aujourd'hui pour l'analyse de données corrélées au moyen de techniques telles que l'analyse longitudinale de données et les équations d'estimation généralisées [*generalized estimating equations* (GEE)]. SUDAAN est l'abréviation de Survey Data Analysis (Analyse des données d'enquêtes). Les procédures statistiques descriptives et analytiques sont dénommées DESCRIPT, CROSTAB et RATIO. Les procédures de modélisation fondée sur la conception comprennent la régression linéaire, la régression logistique (y compris la régression multinominale), la régression log-linéaire et l'analyse de survie.

88. SUDAAN 8.0 est programmé en langage C, avec instructions de commande fournies par l'utilisateur semblables à celles de SAS. Les ensembles de données d'entrée peuvent être des fichiers SAS, SPSS ou ASCII. SUDAAN peut fonctionner de façon autonome (« standalone SUDAAN ») ou combinée avec SAS (« SAS-callable SUDAAN »). Les utilisateurs de SAS préfèrent généralement la deuxième formule.

89. SUDAAN est le seul logiciel d'enquête par sondage qui utilise les deux approches les plus courantes de l'estimation de la variance : la linéarisation par série de Taylor et les méthodes de réplification. Cette dernière approche de SUDAAN comprend la réplification équilibrée répétée [*balanced repeated replication* (BRR)], avec ou sans coefficient d'ajustement de Fay, et les méthodes « jackknife ». Toutes les méthodes de réplification utilisées avec SUDAAN supposent la description courante du plan d'échantillonnage mentionnée précédemment (*WR*). Si l'on utilise la réplification BRR pour l'estimation de la variance, les pondérations des répliqués BRR doivent être données avec l'ensemble de données d'entrée; SUDAAN ne produit pas de pondérations pour les répliqués BRR. SUDAAN produit des pondérations de répliqués pour la méthode « jackknife » qui supprime une UPE, ou accepte des pondérations de répliqués avec l'ensemble de données d'entrée utilisé pour cette méthode ou certaines de ses variantes.

90. On décrit la conception de l'enquête par sondage pour SUDAAN dans trois instructions : *a*) en choisissant une option pour le mot clé DESIGN dans l'instruction PROC; *b*) en spécifiant les variables de stratification et de mise en grappe dans l'instruction NEST; et *c*) en spécifiant la variable de pondération de l'échantillon dans l'instruction WEIGHT. L'ensemble de données de SUDAAN doit être classé selon toutes les variables qui figurent dans l'instruction NEST, qui sont généralement la variable de stratification de la première phase, puis la variable UPE de chaque strate.

91. À la différence des autres logiciels pouvant servir aux enquêtes par sondage, la deuxième phase et les phases suivantes de l'échantillonnage et la stratification d'un échantillonnage à phases multiples peuvent être décrites pour l'estimation de la variance par SUDAAN, ce qui évite d'avoir toujours à utiliser la description courante du plan d'échantillonnage *WR*. En outre, SUDAAN offre l'ample possibilité d'incorporer dans l'estimation de la variance les termes (*fpc*) de correction de la population finie aux phases multiples de l'échantillonnage sans remplacement. Le manuel SUDAAN, qui existe sur papier et en fichier pdf, donne plusieurs exemples de la façon de décrire les plans d'échantillonnage (voir chapitre III).

92. Le plan d'échantillonnage par défaut de SUDAAN est le plan *WR* défini ci-dessus, que ce soit avec la linéarisation par série de Taylor, BRR ou « jackknife ». L'utilisation de la syntaxe DESIGN = *WR* dans l'instruction PROC implique non seulement l'approche UCVE et l'échantillonnage de première phase avec ou sans remplacement mais avec de petites fractions d'échantillonnage de première phase, mais aussi l'utilisation de la linéarisation par série de Taylor. Avec DESIGN = *WR*, l'instruction NEST comprend une variable de justification ou plusieurs (généralement une seule) et une variable UPE. Si l'option DESIGN = ne figure pas dans l'instruction PROC, SUDAAN suppose que DESIGN = *WR*.

93. La syntaxe DESIGN = BRR de SUDAAN implique la description *WR* du plan d'échantillonnage courant (comme on l'a vu précédemment) avec réplification répétée équilibrée pour l'estimation de la variance. Les variables de pondération des répliqués BRR doivent figurer dans l'ensemble des données d'entrée, et l'instruction REPWGT du programme SUDAAN donne les noms des variables pour les variables de pondération des répliqués.

94. La syntaxe DESIGN = JACKKNIFE de SUDAAN, en l'absence d'instructions JACKWGTS et JACKMULT, implique la description WR du plan d'échantillonnage courant avec estimation de la variance par la technique « delete on jackknife ». La syntaxe DESIGN = JACKKNIFE de SUDAAN, avec l'instruction JACKWGTS implique la description WR du plan d'échantillonnage courant avec les pondérations « jackknife » fournies à SUDAAN comme variables dans l'ensemble de données d'entrée.

95. La conception de l'enquête par sondage pour l'enquête du Burundi et la spécification de la linéarisation par série de Taylor pour l'estimation de la variance sont décrites à SUDAAN de la façon suivante :

```
PROC ..... DESIGN = WR .....
NEST PSTR PPSU
WEIGHT W2
```

96. La conception de l'enquête par sondage pour l'enquête du Burundi et la spécification de BRR (réplication répétée équilibrée) pour l'estimation de la variance sont décrites à SUDAAN de la façon suivante :

```
PROC ..... DESIGN = BRR .....
WEIGHT W2
REPWGT REPLWT01-REPLWT32
```

Noter ci-dessus que l'instruction REPWGT identifie les variables de pondération des répliqués inclus dans l'ensemble de données d'entrée. Ces 32 variables de pondération des répliqués sont fondées sur les 30 pseudo-strates, avec 2 UPE par pseudo-strate, et ont été obtenues au moyen de WES VAR. Noter également que l'instruction NEST est absente lorsque l'on utilise BRR; SUDAAN n'a pas besoin de connaître les variables de stratification et les variables des UPE car elle n'utilise que les variables de pondération des répliqués pour l'estimation de la variance.

2. DESCRIPT

97. La procédure DESCRIPT estime les totaux et les moyennes de population pour les variables continues ainsi que les totaux et les moyennes de population pour les variables catégoriques. L'instruction VAR donne la liste des variables (dépendantes) à analyser. Pour un programme DESCRIPT donné, toutes les variables figurant sur l'instruction VAR doivent être continues ou elles doivent toutes être catégoriques. Si les variables catégoriques figurent sur l'instruction VAR, l'instruction CATLEVEL doit aussi être utilisée pour indiquer pour quels niveaux de chaque variable catégorique des estimations sont désirées. Par exemple, les deux instructions ci-dessous estiment les pourcentages de la population étudiée au Burundi qui sont séropositifs et non séropositifs [à supposer que IMMUNE est codé 1, 2 ou (point) pour les données manquantes] :

```
VAR IMMUNE IMMUNE
CATLEVEL 1 2
```

98. On obtient les estimations pour les domaines en utilisant l'instruction TABLES qui contient une ou plusieurs variables catégoriques. On peut comparer les domaines entre eux par contrastes linéaires en utilisant les instructions CONTRAST, PAIRWISE ou DIFFVAR. On peut estimer des taux normalisés et des moyennes, par exemple la prévalence d'une maladie ajustée pour tenir compte de l'âge, en utilisant les instructions STDVAR et STDWGT. On peut évaluer les tendances linéaires et de ni-

veau supérieur (quadratique, etc.) des moyennes ou des pourcentages à travers les niveaux d'une variable catégorique en utilisant l'instruction POLY (POLYNOMIALE); SUDAAN utilise les contrastes linéaires polynomiaux orthogonaux pour ces analyses.

99. Toutes les variables figurant sur une instruction TABLES, CONTRAST, PAIRWISE, DIFFVAR, STDVAR ou POLY doivent figurer sur une instruction SUBGROUP, et une instruction LEVELS indique la valeur codée la plus élevée dans l'analyse de chaque variable catégorique dans l'instruction SUBGROUP.

100. L'instruction SUBPOPN de SUDAAN, qui peut être utilisée dans toutes les PROCS, limite l'analyse à une sous-catégorie de population, par exemple seulement aux femmes âgées. Utiliser l'instruction SUBPOPN avec l'ensemble complet de données de l'enquête par sondage au lieu de restreindre l'ensemble de données d'entrée à la population étudiée avant d'utiliser SUDAAN, car une telle procédure pourrait engendrer des estimations incorrectes des erreurs types dans la mesure où certaines UPE de l'échantillon pourraient ne pas figurer dans l'ensemble de données ainsi restreint.

3. CROSSTAB

101. La procédure CROSSTAB ne convient que pour les variables catégoriques. L'instruction TABLES de CROSSTAB indique le tableau à une, deux ou plusieurs entrées pour lequel les pourcentages et les totaux de population sont estimés. Des instructions correspondantes SUBGROUP et LEVELS sont nécessaires pour toutes les variables figurant sur l'instruction TABLES.

102. L'instruction TEST de CROSSTAB demande des tests de chi-carré pour tester l'hypothèse nulle selon laquelle deux variables catégoriques sont statistiquement indépendantes. Un test de chi-carré est fondé sur le test type Pearson (CHISQ), qui utilise des calculs « observés moins attendus » pour les estimations des totaux de population. L'autre test de chi-carré est fondé sur les estimations des odd-ratios de la population (LLCHISQ). On estime les odds-ratios et les risques relatifs (ratios de prévalence réelle), avec intervalles de confiance, pour les tableaux 2×2 en utilisant RISK = ALL sur l'instruction PRINT. Enfin, on dispose d'un test de Cochran-Mantel-Haenszel (l'instruction TEST utilise CMH) pour évaluer l'indépendance statistique de deux variables tout en contrôlant (en « stratifiant ») sur une troisième variable.

4. Exemples numériques

103. La section III de l'annexe sur le CD-ROM illustre l'utilisation de CROSSTAB et DESCRIPT pour étudier les trois exemples énumérés au paragraphe 60, utilisant SAS-CALLABLE SUDAAN (SAS Version 8.2 et SUDAAN Version 8.0). On utilise à la fois la linéarisation par série de Taylor et BRR (la réplication répétée équilibrée) pour estimer la variance. L'examen des programmes SUDAAN annotés (par les utilisateurs) et des résultats de SUDAAN annotés devraient aider les lecteurs à écrire leurs propres programmes SUDAAN et à en interpréter les résultats. Seules certaines analyses sélectionnées de SUDAAN examinées aux TABLES 1, 3, 4, 5 et 6 sont incluses et annotées à la section III de l'annexe.

104. Le tableau XXI.1, rang 4, récapitule les résultats de CROSSTAB et DESCRIPT dans la section III (annexe) pour estimer le pourcentage et le nombre de femmes de la population étudiée au Burundi qui sont séropositives, avec estimation de l'erreur type. Les résultats de CROSSTAB et DESCRIPT obtenus avec SUDAAN sont identiques pour une méthode donnée d'estimation de la variance (comme prévu), et

les résultats de la série de Taylor et de BRR sont identiques (ce qui n'est pas toujours le cas). Noter que CROSSTAB et DESCRIPT ne calculent pas les intervalles de confiance pour les estimations des pourcentages et des totaux de population.

105. Le tableau XXI.3, rang 2, de l'annexe, section VII (CD-ROM) montre que l'on obtient des résultats identiques avec CROSSTAB et DESCRIPT (que ce soit avec la série de Taylor ou avec BRR) pour l'estimation du pourcentage de femmes séropositives, dans chacun des deux domaines des femmes rurales et urbaines. Les résultats de SUDAAN CROSSTAB et DESCRIPT concordent avec SAS SURVEYMEANS.

106. Le tableau XXI.4, rang 2, de l'annexe, section VII (CD-ROM), résume les résultats de DESCRIPT (avec la série de Taylor et BRR) qui utilise un contraste linéaire pour comparer les femmes rurales aux femmes urbaines pour ce qui est du pourcentage de cas de séropositivité. Il y a une différence négligeable dans l'erreur type estimée avec la série de Taylor et avec BRR. La conclusion est la suivante : il y a une différence de séropositivité entre les femmes urbaines et rurales de la population étudiée du Burundi; la séropositivité est plus élevée chez les femmes urbaines. On notera que les résultats du contraste linéaire de DESCRIPT concordent avec ceux obtenus quand on utilise SAS SURVEYREG pour comparer les deux domaines.

107. Le tableau XXI.5, rangs 1 et 2, de l'annexe, section VII (CD-ROM), indique les résultats des deux tests de chi-carré différents qui peuvent être effectués avec CROSSTAB : Pearson (CHISQ) et log-linéaire (LLCHISQ). Les résultats obtenus avec la série de Taylor et BRR sont identiques. La séropositivité estimée est sensiblement plus élevée chez les femmes urbaines que chez les femmes rurales (avec CHISQ), et les odds estimés de séropositivité sont sensiblement plus élevés pour les femmes urbaines que pour les femmes rurales (avec LLCHISQ).

108. Le tableau XXI.6, rang 1, de l'annexe, section VII (CD-ROM), indique l'odds-ratio estimé (0,393) et le ratio de prévalence (0,797) pour la séropositivité (rurale/urbaine), chacun avec un intervalle de confiance de 0,95 %. La série de Taylor et BRR présentent des différences négligeables vers la limite supérieure de l'intervalle de confiance de 95 % sur l'odds-ratio. L'odds-ratio estimé et le ratio de prévalence diffèrent par leur ampleur parce que la prévalence de la séropositivité n'est pas faible.

5. Avantages/inconvénients/coût

109. SUDAAN est un logiciel complet d'enquête par sondage (et de données corrélées) avec capacité d'analyse aussi bien descriptive que de modélisation. Il offre une vaste capacité d'estimer et de tester les matrices de contraste spécifiées par l'utilisateur sur les paramètres de population, y compris les coefficients de régression. Il peut être utilisé aussi bien sur ordinateur central que sur ordinateur individuel. Les utilisateurs de SAS ont intérêt à apprendre SUDAAN, car la syntaxe est similaire à celle de SAS. Toutefois, une partie de la syntaxe de SUDAAN est ésothérique et nécessite peut-être plus de temps d'étude que les autres logiciels.

110. Comparé à celui des autres logiciels examinés dans ce chapitre, le coût de SUDAAN est élevé, surtout s'il est utilisé comme SUDAAN callable dans SAS, parce que dans ce cas, il faut aussi SAS. Un appui technique est dispensé aux utilisateurs ayant une licence d'utilisation. Le Manuel d'utilisation de SUDAAN Version 8.0, principalement en tant qu'ouvrage de référence plutôt que pour apprendre SUDAAN, offre plusieurs exemples annotés d'analyses avec des données de NHANES-III (*National Health and Nutrition Examination Survey-III*) [Enquête nationale sur la santé et la nutrition] qui peuvent être utiles pour apprendre comment utiliser SUDAAN.

111. SUDAAN est le seul logiciel illustré ici qui offre deux importantes approches de l'estimation de la variance, la linéarisation par série de Taylor et les méthodes de réplification. Cependant, SUDAAN n'établit pas de pondérations de répliqués pour la réplification répétée équilibrée, ce qui oblige l'utilisateur à fournir ces pondérations. SUDAAN établit des pondérations de répliqués pour la procédure « jackknife delete one » et accepte aussi les pondérations des répliqués de « jackknife » si elles figurent dans l'ensemble de données d'entrée.

112. SUDAAN est le seul logiciel examiné ici qui offre une large capacité de description de plusieurs phases d'échantillonnage, de stratification et des termes *fpc* à incorporer dans l'estimation de la variance. Par ailleurs, il offre plusieurs définitions différentes pour les calculs des effets de conception pour permettre d'exclure de l'effet de conception les effets du suréchantillonnage et/ou de pondérations inégales.

113. L'entrée de données ASCII dans SUDAAN est laborieuse et rend préférables les deux autres options d'entrée de données, à savoir un ensemble de données SAS ou SPSS. L'entrée d'un ensemble de données SAS dans SUDAAN doit être la Version 6.04 de SAS ou un fichier de transport SAS. SUDAAN appelable dans SAS peut lire n'importe quel ensemble de données que peut lire SAS. La sortie de SUDAAN peut être sauvegardée électroniquement sur un format de fichier de données SAS pour pouvoir être utilisé dans SAS ou avec un logiciel de tableur tel qu'EXCEL. SUDAAN a une capacité très limitée de recoder les variables et n'a aucune capacité de gestion de données. Il est donc prudent de recoder et de former les nouvelles variables dans SAS ou SPSS (selon le type d'ensemble de données d'entrée) avant d'utiliser SUDAAN.

I. Procédures d'enquête par sondage dans STATA 7.0

1. Description générale de STATA

114. STATA est un logiciel général de statistique, créé en 1995, qui ajoute une vaste capacité d'analyse des données d'enquêtes par sondage. STATA 7.0 est présenté ici. La version 8.0 a été lancée en 2003. Seule la linéarisation par série de Taylor est utilisée pour estimer la variance. La description du plan d'échantillonnage courant *WR* est la version par défaut. STATA peut inclure les termes *fpc* dans l'estimation de la variance pour une phase unique sans plans d'échantillonnage de remplacement (échantillonnage aléatoire simple et échantillonnage aléatoire stratifié) et pour l'échantillonnage en grappe à une phase sans remplacement (stratifié ou non) où l'on utilise l'égalité de probabilité d'échantillonnage pour les grappes (UPE) à l'intérieur d'une strate et où tous les éléments d'une UPE sélectionnée sont inclus dans l'échantillon.

115. L'ampleur des analyses des données d'enquêtes par sondage effectuées par STATA se compare favorablement à celle de SUDAAN, avec une capacité statistique mathématique pour matrices de contraste spécifiées par l'utilisateur sur les paramètres de population, y compris les coefficients de régression. STATA fonctionne de façon interactive sur commandes brèves et simples, ce qui le rend relativement facile à apprendre. Toutefois, les programmes conçus par les utilisateurs peuvent éventuellement être soumis par séries. STATA ne répond qu'à des commandes en minuscules. STATA alloue un volume de mémoire par défaut sur lequel il charge une copie de l'ensemble de données d'entrée. Si cette mémoire est insuffisante pour les gros ensembles de données, elle peut être augmentée par utilisation de la commande de réglage de mémoire.

116. Les commandes d'enquête par sondage de STATA commencent par le nom *svy* (survey = enquête). Des commandes descriptives sont à utiliser pour estimer une moyenne de population (*svymean*), un total de population (*svytotal*), une pro-

portion de population (`svyprop`) et des pourcentages et des totaux dans des tableaux à deux entrées (`svytab`). Les intervalles de confiance sur les proportions de population de `svytab` utilisent une transformation logit, de sorte que les limites estimatives inférieure et supérieure sont données à (0, 1). On dispose de huit tests de chi-carré différents pour les données d'enquêtes par sondage sur tableaux à deux entrées dans `svytab`. Les procédures de modélisation utilisables comprennent la régression linéaire, la régression logistique (y compris la régression multinominale avec une variable nominale ou ordonnée), la régression de Poisson et les modèles probit.

117. La commande `svyset` sert à spécifier le plan d'échantillonnage à STATA. Pour décrire le plan d'échantillonnage courant `WR` (défaut), trois mots clés sont tapés de façon interactive dans STATA pour la commande `svyset`. Le mot clé `strata` précède le nom de la variable de stratification; le mot clé `psu` (UPE) précède le nom de la variable UPE, et le mot clé `pweight` précède le nom de la variable de pondération d'échantillonnage. Ainsi, le plan d'échantillonnage pour l'enquête effectuée au Burundi est décrit dans STATA V7 comme

```
svyset strata pstra
svyset psu ppsu
svyset pweight w2
```

118. Comme on l'a indiqué précédemment pour les procédures d'enquête par sondage dans SAS, l'omission du mot clé `strata` dans STATA implique l'absence de stratification des UPE avant l'échantillonnage de première phase. L'omission du mot clé `psu` (UPE) implique un échantillonnage à une phase des éléments et pas de mise en grappe des éléments sélectionnés. L'omission du mot clé `pweight` implique la présence d'éléments de pondération égale, avec une valeur par défaut de 1,0 pour la variable de pondération. La syntaxe pour la commande `svyset` est révisée dans STATA V8.

119. La commande `svydes` ordonne à STATA de sortir les variables de conception de l'enquête qu'il a jointes à l'ensemble de données (à partir des commandes `svyset`) et de récapituler le nombre de strates, le nombre d'UPE par strate et le nombre moyen d'observations par UPE à l'intérieur de chaque strate. C'est là un résumé très utile des caractéristiques de la conception de l'enquête par sondage.

2. `Svymean`, `svyprop`, `svytotal`, `svylc`

120. La commande `svymean` estime une moyenne de population, soit pour une variable continue soit pour une variable indicatrice codée 1 ou 0 (c'est-à-dire une proportion estimative de la population). Les options de sortie comprennent l'erreur type estimative, le coefficient estimatif de variation, l'effet de conception et l'intervalle de confiance sur le paramètre de population.

121. La commande `svyprop` est pour les données catégoriques : elle estime la proportion de la population qui se situe à chaque niveau de la variable catégorique, ainsi que l'erreur type estimative. Cette commande offre un moins grand nombre d'options de sortie que `svymean`.

122. La commande `svytotal` estime un total de population pour une variable continue ou une variable indicatrice (0,1), avec erreur type estimative, coefficient estimatif de variation, effet de conception et intervalle de confiance.

123. Chacune des trois commandes ci-dessus peut servir à estimer des paramètres de population pour les domaines en utilisant l'option `by` sur la ligne de commande, par exemple, `by (stra)` ou `by (urb_rur)` pour analyser les deux domaines des femmes rurales et urbaines au Burundi. STATA utilise des formules correctes d'esti-

mation de la variance pour les domaines avec l'instruction `by` dans ses commandes `svy`.

124. En outre, chacune des trois commandes ci-dessus peut être utilisée avec une option `subprop` sur la ligne de commande pour effectuer une estimation des paramètres de population d'une catégorie donnée de population, par exemple les femmes âgées. Ne pas utiliser l'instruction `if` de STATA pour les analyses de fragments de la population parce que les variances estimatives peuvent être incorrectes; utiliser l'option `subprop`.

125. La commande `svylc` estime les combinaisons linéaires spécifiées par l'utilisateur des moyennes, des proportions ou des totaux des domaines, ainsi que l'erreur type estimative, le test-t, la valeur-p et l'intervalle de confiance. Cette commande peut être utilisée pour comparer des domaines entre eux. Dans V8.0, la commande `svylc` est remplacée par `lincom`. La commande `svylc` reste utilisable avec V8.0 mais elle n'est plus documentée.

3. Svytab

126. La commande `svytab` de STATA est à utiliser pour les tableaux à deux entrées. Elle permet d'estimer les pourcentages de population (rang, colonne ou total) avec des erreurs types estimatives, les totaux de population pour les cellules des tableaux avec des erreurs types estimatives, et les intervalles de confiance. Un `logit transform` permet d'obtenir les intervalles de confiance sur les proportions de population, de sorte que les limites estimatives inférieure et supérieure sont tenues de se situer dans cet intervalle (0, 1). On dispose de huit différents tests de chi-carré pour tester l'hypothèse nulle de l'indépendance statistique des deux variables catégoriques du tableau. La commande `subpop` est à utiliser avec `svytab`.

4. Exemples numériques

127. La section IV de l'annexe (CD-ROM) décrit l'utilisation des commandes STATA avec les trois exemples indiqués au paragraphe 60. Chaque exemple est un fichier `log` de la session interactive avec STATA. L'examen du `log STATA` annoté (commandes d'utilisateur et résultats de STATA) devrait aider les lecteurs pour l'utilisation des commandes d'enquête par sondage de STATA et l'interprétation des résultats.

128. Les commandes `svymeans` et `svytotal` ont été utilisées avec la variable indicatrice `BLOOD` (1 = séropositive, 0 = séronégative). Le tableau XXI.1 (rangs 5 et 6) indique le nombre et le pourcentage estimatifs de femmes séropositives avec des intervalles de confiance. Les calculs de STATA concordent avec ceux de SAS SURVEYMEANS et de SUDAAN DESCRIPT et CROSSTAB.

129. Le tableau XXI.3 (rang 5) de l'annexe, section VII (CD-ROM), indique le pourcentage estimatif de femmes séropositives selon leur lieu de résidence (rural/urbain). Les estimations ponctuelles STATA `svytab` et les erreurs types estimatives concordent avec celles de SAS SURVEYMEANS et de SUDAAN DESCRIPT et CROSSTAB. Toutefois, les intervalles de confiance pour les domaines diffèrent légèrement entre STATA `svytab` et SAS SURVEYMEANS parce que STATA `svytab` utilise un `logit transform` pour obtenir les intervalles de confiance.

130. Le tableau XXI.4 (rang 3) de l'annexe, section VII (CD-ROM), présente les résultats de STATA `svylc` pour le contraste linéaire qui compare les femmes rurales et urbaines pour ce qui est du pourcentage de femmes séropositives, indiquant une différence sensible entre les deux domaines. Les résultats de STATA concordent avec

ceux de SUDAAN DESCRIPT et avec l'utilisation de SAS SURVEYREG pour comparer les domaines.

131. Le tableau XXI.5 (rangs 3 à 5) de l'annexe, section VII (CD-ROM), présente les résultats de STATA svytab pour trois tests de chi-carré de l'hypothèse nulle selon laquelle la séropositivité est statistiquement indépendante du lieu d'habitation (rural/urbain). Ces trois tests ont tous des valeurs-p similaires (et faibles). Le test de chi-carré par défaut pour STATA svytab (rang 3) est un test de chi-carré de type Pearson proposé par Rao et Scott (1981; 1984) avec une correction de second ordre. Les deux autres tests de chi-carré de svytab (rangs 4 et 5) sont les mêmes que dans SUDAAN CROSSTAB, et STATA et SUDAAN produisent les mêmes calculs pour ces deux tests.

132. Comme la commande svytab de STATA ne produit pas d'odds-ratios ou de ratios de prévalence, la commande svylogit a été utilisée pour estimer l'odds-ratio (urbain/rural) pour la séropositivité. L'odds-ratio de STATA, avec intervalle de confiance, figure au tableau XXI.6 (rang 2) de l'annexe, section VII (CD-ROM). La commande STATA svylogit donne les mêmes calculs que SUDAAN CROSSTAB pour l'estimation ponctuelle et l'intervalle de confiance.

5. Avantages/inconvénients/coût

133. STATA est un logiciel complet d'analyse générale de statistiques qui offre une vaste capacité d'analyse des données d'enquêtes par sondage, y compris des procédures de modélisation descriptive et de modélisation fondée sur la conception. Il offre de nombreuses procédures de modélisation pour les données d'enquêtes par sondage. STATA a été très bien accueilli en tant que logiciel de statistique; il est relativement facile à apprendre et dispose d'un groupe d'utilisateurs actifs. Comparé aux autres logiciels examinés dans ce chapitre, il est d'un coût modéré.

134. STATA accepte les matrices de contraste des paramètres estimatifs de population définies par les utilisateurs, y compris les coefficients de régression, pour ceux qui désirent tester leurs propres hypothèses ou estimer des combinaisons de paramètres de population. En général, il offre une grande souplesse dans la conduite d'analyses statistiques à ceux qui possèdent les connaissances mathématiques statistiques requises.

135. STATA utilise la linéarisation par série de Taylor et est limité à la description du plan d'échantillonnage courant *WR*. Toutefois, il peut inclure dans l'estimation de la variance pour l'échantillonnage sans remplacement les termes *fpc* pour l'échantillonnage à une phase d'éléments et pour l'échantillonnage en grappe à une phase. Il est quelque peu difficile, mais possible, d'extraire des résultats analytiques de STATA (par exemple, les tailles des échantillons non pondérés, les estimations ponctuelles et les erreurs types) pour les transposer à d'autres formats de données.

J. Procédures d'enquête par sondage dans Epi-Info 6.04d et Epi-Info 2002

1. Aperçu général d'Epi-Info

136. Epi-Info a été conçu au terme de nombreuses années de travail par les Centers for Disease Control and Prevention (CDC) et l'Organisation mondiale de la santé (OMS). Ce logiciel est offert gratuitement sur simple téléchargement à partir du site Internet des CDC : <http://www.cdc.gov/epiinfo/>.

137. Il existe deux versions d'Epi-Info : la dernière version DOS Epi-Info Version 6.04d et la toute récente version Windows Epi-Info 2002.

138. Epi-Info permet, entre autres, l'élaboration d'un questionnaire ou d'un formulaire de collecte de données de recherche, l'entrée de données adaptées aux besoins du client, l'analyse des données et le traitement de texte. Ses capacités analytiques et statistiques s'adressent aux épidémiologistes du monde entier. Les résultats analytiques des analyses d'Epi-Info peuvent être transmis sur écran, sur imprimante ou sur fichier électronique.

139. Les deux versions d'Epi-Info (DOS et Windows) peuvent effectuer des analyses descriptives de base de données d'enquêtes par sondage complexes. Elles n'utilisent que la description *WR* du plan d'échantillonnage courant. Les données d'entrée doivent être triées selon deux des trois variables de conception de l'enquête : la variable de stratification *STRATVAR* et la variable PSU (UPE) *PSUVAR* de la strate. Epi-Info n'incorpore pas les termes *fpc* dans l'estimation de la variance. En outre, elle ne donne pas d'estimation des totaux de population. Elle utilise la linéarisation par série de Taylor pour estimer la variance.

140. La capacité analytique d'Epi-Info pour les données d'enquêtes par sondage complexes a été conçue à l'origine pour le Système de surveillance du facteur de risque du comportement [*Behavioral Risk Factor Surveillance System* (BRFSS)], programme annuel d'enquête par sondage sur la santé parrainé par les CDC pour les États-Unis d'Amérique (Brogan, 1998 et sous presse) et pour la méthodologie d'échantillonnage en grappe de l'OMS utilisée dans le monde entier dans le cadre du Programme élargi de vaccination [*Expanded Programme on Immunization* (EPI)] pour estimer la couverture de vaccination des enfants (Brogan *et al.*, 1994). Toutefois, les procédures d'enquête par sondage d'Epi-Info peuvent être utilisées pour n'importe quelle enquête par sondage complexe pouvant être décrite par la description *WR* du plan d'échantillonnage courant.

2. Epi-Info Version 6.04d (DOS), module CSAMPLE

141. La version DOS d'Epi-Info est le fruit d'un effort commun des CDC et de l'OMS. L'entrée des données d'Epi-Info 6.04d est un fichier dBase ou un fichier ASCII qu'Epi-Info convertit ensuite en un fichier de données Epi-Info *.rec, par exemple DBMS-COPY (<http://www.dataflux.com/conceptual/>). Epi-Info 6.04d utilise un programme interactif et ne peut être utilisé en mode batch. La version DOS peut avoir la préférence sur la version Windows chez ceux qui ont des ordinateurs ou des systèmes d'exploitation plus anciens et/ou qui ne disposent que d'une capacité de mémoire limitée sur leur disque dur.

142. Le module CSAMPLE de la Version 6.04d d'Epi-Info effectue des analyses de données d'enquêtes par sondage complexes. CSAMPLE permet d'estimer une moyenne de population (pour une variable continue ou une variable indicatrice codée 1/0) ou un pourcentage de population (pour une variable catégorique), ainsi que l'erreur type, les intervalles de confiance et l'effet de conception. Ces estimations sont également données pour les domaines formés par les niveaux d'une variable catégorique. En outre, CSAMPLE permet d'estimer la différence entre des moyennes de domaines ou des pourcentages de domaines, avec l'erreur type estimative correspondante et un intervalle de confiance sur la différence de population. CSAMPLE estime l'odds-ratio et le ratio de risque pour les tableaux 2×2 . Il convient de noter que CSAMPLE ne donne pas d'estimations des totaux de population.

143. Lorsque le module CSAMPLE est ouvert dans Epi-Info 6.04d, un écran d'entrée de données apparaît sur lequel l'utilisateur spécifie les variables à utiliser pour l'analyse. L'utilisateur choisit une variable pour chacune des trois cases de conception de l'enquête : STRATA (la variable de stratification), PSU (la variable UPE ou grappe) et WEIGHT (la variable de pondération de l'échantillonnage). Pour l'enquête effectuée au Burundi, les variables spécifiées à Epi-Info étaient les suivantes :

```
STRATA  PSTRA
PSU     PPSU
WEIGHT  W2
```

144. L'utilisateur spécifie la variable d'analyse (ou variable dépendante) dans la case appelée MAIN. Cette variable peut être continue, par exemple *IUML*, ou catégorique, par exemple *IMMUNE*. S'il désire une estimation de la moyenne de population pour la variable continue (ou censée être continue) spécifiée dans la case MAIN, l'utilisateur clique sur l'option MEANS. S'il désire une estimation des pourcentages de population pour la variable catégorique spécifiée dans la case MAIN, l'utilisateur clique sur l'option TABLE.

145. S'il désire des estimations de moyenne ou des pourcentages de domaines, l'utilisateur spécifie la variable définissant les domaines dans la case appelée CROSSTAB et la variable d'analyse dans la case MAIN.

146. Par ailleurs, CSAMPLE permet d'estimer la différence entre deux domaines sur la moyenne d'une variable d'analyse. L'utilisateur peut spécifier deux niveaux de la variable CROSSTAB qui définit les deux domaines à comparer entre eux.

3. Epi-Info 2002 (Windows)

147. Epi-Info 2002 est une application de Windows conçue par les CDC. L'entrée des données pour l'analyse par Epi-Info 2002 se fait par fichier 1997 d'accès Microsoft (*.mdb) ou sur un fichier dBase. Epi-Info 2002 peut aussi lire les fichiers *.rec préparés pour les versions DOS d'Epi-Info. Le logiciel fonctionne de façon interactive mais peut fonctionner en mode batch.

148. Epi-Info 2002 a trois procédures d'échantillonnage complexe situées dans la section Analyse de données sous le titre Statistiques avancées. Les fréquences d'échantillonnage complexe estiment une distribution en pourcentage pour un tableau à une entrée d'une variable catégorique avec estimation de l'erreur type et des intervalles de confiance. Les tableaux d'échantillonnage complexe estiment les pourcentages en rang et en colonne pour un tableau de variables catégoriques à deux entrées [dénommées exposition (rang) et résultats (colonne)] avec estimation de l'erreur type et des intervalles de confiance pour les pourcentages de rang. S'il s'agit d'un tableau 2 × 2, la procédure estime également l'odds-ratio et le ratio de risque, avec intervalles de confiance. La moyenne d'échantillonnage complexe estime la moyenne d'une variable continue, avec estimation de l'erreur type et de l'intervalle de confiance, y compris l'estimation de la moyenne pour les domaines formés par une variable catégorique. Si la variable domaine se situe à deux niveaux, la différence entre la moyenne du domaine est aussi estimée, de même que l'erreur type et l'intervalle de confiance.

149. Avec ces trois procédures d'échantillonnage complexe, les variables de conception de l'enquête sont indiquées dans trois cases dénommées Weight, PSU et Stratify By (la variable de stratification de l'enquête par sondage). Pour obtenir des estimations des erreurs types et des intervalles de confiance comme sortie, cliquer deux fois sur OPTIONS SET, puis choisir Statistics = Advanced.

4. Exemples numériques

150. La section V de l'annexe (CD-ROM) décrit l'utilisation de CSAMPLE avec Epi-Info 6.04d pour analyser les trois exemples du paragraphe 60. Chaque exemple analysé contient les résultats d'Epi-Info annotés avec commentaires. L'examen de ces résultats annotés devrait aider le lecteur à interpréter les résultats de CSAMPLE.

151. Le tableau XXI.1 (rang 7) donne l'estimation d'Epi-Info 6.04d du pourcentage de femmes qui sont séropositives. L'estimation ponctuelle et l'erreur type estimative d'Epi-Info concordent avec SAS SURVEYMEANS, STATA svymean, SUDAAN DESCRIPT et CROSSTAB. L'intervalle de confiance de 95 % sur la prévalence de la séropositivité est plus étroit que celui donné par SAS SURVEYMEANS et STATA svymean. Cela est dû au fait qu'Epi-Info utilise $z = 1,96$ pour son calcul de l'intervalle de confiance au lieu de la valeur Student-t de 2,042 avec 30 df, les degrés de liberté du dénominateur pour l'enquête effectuée au Burundi [nombre d'UPE (60) moins nombre de pseudo-strates (30)].

152. Le tableau XXI.3 (rang 4) de la section VII de l'annexe (CD-ROM) donne les estimations d'Epi-Info de la prévalence de la séropositivité selon la résidence rurale/urbaine. Les estimations ponctuelles et les erreurs types estimatives concordent avec SAS SURVEYMEANS, STATA svytab, SUDAAN DESCRIPT et CROSSTAB. Les intervalles de confiance d'Epi-Info pour les domaines sont plus étroits que ceux de SAS SURVEYMEANS et STATA svytab parce qu'Epi-Info utilise $z = 1,96$.

153. Le tableau XXI.4 (rang 4) de la section VII de l'annexe (CD-ROM) donne le résultat du contraste linéaire d'Epi-Info qui compare la prévalence de la séropositivité entre femmes rurales et urbaines. La valeur estimative du contraste ($-16,99$) concorde avec SAS SURVEYREG, avec SUDAAN DESCRIPT et avec STATA svylc. Epi-Info ne donne pas l'estimation de l'erreur type de la différence estimative, et l'intervalle de confiance de 95 % que donne Epi-Info sur la valeur du contraste est erronée.

154. Le tableau XXI.6 (rang 3) de la section VII de l'annexe (CD-ROM) donne l'estimation de l'odds-ratio d'Epi-Info (urbain/rural) et l'estimation du ratio de prévalence de la séropositivité, avec un intervalle de confiance de 95 %. Les estimations ponctuelles d'Epi-Info coïncident exactement avec SUDAAN CROSSTAB et avec STATA svylogit, et les intervalles de confiance d'Epi-Info sont étroitement en accord avec SUDAAN et STATA.

5. Avantages/inconvénients/coût

155. Le gros avantage d'Epi-Info est son coût : il peut être téléchargé du site web des CDC. De plus, il est disponible à la fois pour les systèmes d'exploitation DOS et WINDOWS, ce qui confère une grande souplesse de choix des matériels et des logiciels avec lesquels l'utiliser. La capacité d'utilisation d'Epi-Info pour des enquêtes par sondage est certainement un attrait pour ceux qui l'utilisent déjà pour d'autres types d'analyses épidémiologiques ou statistiques.

156. Epi-Info n'utilise que la linéarisation par série de Taylor et est limité à la description du plan d'échantillonnage courant *WR*. Le module CSAMPLE de la version DOS et son homologue de la version Windows (trois procédures d'Advanced Statistics) conviennent pour des statistiques descriptives de base pour données d'enquêtes par sondage complexes, notamment pour l'estimation de moyennes de population ou de pourcentages de la population entière et pour les domaines, ainsi que pour les comparaisons de domaines. Epi-Info n'offre pas de capacité d'enquête par sondage pour estimer les totaux de population, pour effectuer des tests de chi-carré, pour in-

corpore les termes fpc (correction de la population finie) dans l'estimation de la variance ou pour des analyses de modélisation fondées sur la conception (par exemple, régression logistique ou régression linéaire).

K. WesVar 4.2

1. Aperçu général de WesVar

157. WesVar est un logiciel d'analyse des données d'enquêtes par sondage. Il utilise les méthodes de réplication (Rust et Rao, 1996) pour l'estimation de la variance : BRR, y compris le facteur Fay facultatif et les trois variantes « jackknife ». WesVar n'a pas la capacité d'utilisation de la linéarisation par série de Taylor. Les conceptions d'enquêtes par sondage qui se prêtent bien à l'utilisation de BRR ont plusieurs strates et exactement deux UPE sélectionnées par strate. Les méthodes « jackknife », comme la linéarisation par série de Taylor, peuvent s'appliquer à une conception avec n'importe quel nombre ($> = 2$) d'UPE sélectionnées par strate.

158. Le plan d'échantillonnage par défaut pour WesVar est le plan d'échantillonnage courant *WR* mentionné précédemment. WesVar peut inclure les facteurs fpc dans l'estimation de la variance, mais seulement pour les techniques « jackknife » et seulement pour l'échantillonnage à une phase d'éléments.

159. WesVar 4.2 peut lire les types d'ensembles de données suivants : PC-SAS pour DOS, SAS transport, SAS (versions 6-8), SPSS, STATA, ASCII, et fichiers conformes à ODBC tels que Microsoft Excel ou Access. Conformément au plan d'échantillonnage courant *WR*, s'il faut construire des variables de pondération, WesVar nécessite les variables de stratification, d'UPE et de pondération pour chaque observation. Toutefois, une fois que les pondérations des répliqués sont sur le fichier, on n'a plus besoin des identificateurs d'UPE et de strate : c'est là un avantage de confidentialité des méthodes de réplication pour les fichiers d'utilisation publique. WesVar est le seul des logiciels examinés qui peut ajuster les pondérations de base de l'enquête pour les cas de non-réponse, la post-stratification et le ratissage. Une fois que la préparation de l'ensemble de données d'entrée est terminée, cet ensemble est sauvegardé comme fichier WesVar (*.var) pour l'analyse des données et toute gestion future des données.

160. WesVar offre un ensemble complet de statistiques descriptives : estimations de moyennes, pourcentages, pourcentiles et totaux de population, avec erreur type estimative, coefficient de variation, intervalle de confiance et effet de conception. L'un des points forts de WesVar, et des méthodes de réplication en général, est qu'il permet d'obtenir des estimations ponctuelles (avec erreur type estimative) de fonctions spécifiées par l'utilisateur de paramètres de population, par exemple de ratios de prévalence. Il peut également effectuer des analyses de régression fondées sur la conception : linéaire, logistique et logistique multinominale.

161. On peut télécharger une Version 4 de WesVar à partir de la page web WES-TAT pour une période d'essai de trente jours. La Version 2 peut être téléchargée de la page web et utilisée gratuitement pendant une période illimitée (voir <http://www.wes-tat.com/wesvar>). WesVar Version 4, comparé à la Version 2, accepte une plus grande variété d'ensembles de données d'entrée, a une meilleure capacité de traitement des fichiers et de gestion des données, ajuste les pondérations des répliqués pour les cas de non-réponse et offre de nombreuses autres options analytiques. L'utilisateur pourrait commencer par WesVar Version 2, puis passer éventuellement à la Version 4.

2. Utilisation de WesVar Version 4.2

162. Avec Windows, l'interaction entre l'utilisateur et WesVar se fait par pop-up de menus sur l'écran. Lorsque l'on ouvre le logiciel WesVar, le premier menu offre quatre options. La première option, le nouveau fichier de données WesVar : 1) lit un ensemble de données d'entrée qui n'est pas un ensemble de données WesVar; 2) crée des pondérations de répliqués ou accepte des pondérations de répliqués qui figurent déjà dans l'ensemble de données d'entrée; 3) recode, transforme, étiquette et formate les variables; 4) effectue une post-stratification, un ratissage et des ajustements pour cas de non-réponse; 5) définit des catégories données de population à analyser; et 6) modifie éventuellement le ddf par défaut, puis sauvegarde l'ensemble de données comme fichier WesVar. La deuxième option ouvre le fichier de données WesVar, lit ce fichier et autorise les six opérations précédentes.

163. La troisième option, New WesVar Notebook, accepte les demandes d'analyse d'un fichier de données WesVar, traite ces demandes, affiche les résultats et sauvegarde les demandes et les résultats dans un « bloc-notes » (« notebook »), selon le système WesVar d'organisation des demandes d'analyse et des résultats obtenus. L'un des types d'analyse d'un ensemble de deux est demandé : tableaux ou régression (linéaire, logistique ou multinominale). Une fois que le choix entre tableaux et régression est fait, il existe plusieurs options pour spécifier l'analyse. La lecture des écrans de menu pour analyser et lire les résultats n'est pas immédiate, mais le Guide de l'utilisateur de WesVar offre plusieurs exemples utiles pour illustrer la lecture du menu et l'organisation des résultats.

164. Si les demandes et les résultats d'une précédente session de WesVar ont été sauvegardés sur un bloc-notes, on peut alors choisir la quatrième option du premier menu : ouvrir le bloc-notes WesVar. De nouvelles demandes d'analyse peuvent être ajoutées à un bloc-notes existant, puis sauvegardées. Toutes les analyses liées à un fichier de données WesVar particulier ou à un projet particulier peuvent être organisées en un ou plusieurs blocs-notes.

165. Il faut spécifier l'une des cinq méthodes de réplification de WesVar pour calculer des pondérations de répliqués ou reconnaître de telles pondérations qui existent déjà dans le fichier de données d'entrée. Ces méthodes de réplification sont :

- a) Réplification répétée équilibrée (BRR), exactement deux UPE sélectionnées par strate;
- b) Méthode de perturbation de Fay (FAY) avec BRR;
- c) « Jackknife delete one » sans stratification explicite (JK1);
- d) « Jackknife » avec exactement deux UPE sélectionnées par strate (JK2);
- e) « Jackknife » avec deux UPE sélectionnées par strate ou plus (JKn).

166. Les appendices A et D du *Guide de l'utilisateur* de WesVar donnent un excellent aperçu de ces cinq méthodes de réplification au moyen d'exemples de la façon de passer de divers plans d'échantillonnage à l'une de ces cinq méthodes.

3. Exemples numériques

167. Comme l'ensemble de données d'entrée de l'enquête effectuée au Burundi ne contenait pas de pondérations de répliqués, il fallait choisir l'une des cinq techniques de réplification disponibles puis demander à WesVar de calculer les pondérations des répliqués. Les variables de l'enquête effectuée au Burundi dont avait besoin WesVar étaient : *PSTRA*, *PPSU* et *W2*. Comme le plan d'échantillonnage de cette enquête est *WR*, avec 30 pseudo-strates et exactement deux UPE sélectionnées par strate, BRR

et JK2 sont les meilleurs choix. BRR a été choisi sans facteur de perturbation de Fay. En outre, aucun ajustement pour cas de non-réponse, aucune post-stratification ou aucun ratissage n'a été effectué par les répliqués car ces ajustements n'avaient pas été effectués sur l'ensemble complet de données lorsqu'a été utilisée la linéarisation par série de Taylor.

168. La section VI (CD-ROM) illustre l'utilisation de WesVar sur les trois exemples du paragraphe 60. Chaque exemple donne les résultats de WesVar 4.1 ou 4.2, mais les écrans des menus d'entrée pour les analyses demandées ne sont pas représentés. L'examen des résultats annotés de WESVAR devraient aider le lecteur à interpréter les résultats de WesVar.

169. Le tableau XXI.1, rang 8, montre que WesVar concorde avec tous les autres logiciels d'enquête par sondage sur le pourcentage estimatif et le nombre estimatif de femmes séropositives (avec erreurs types). Les intervalles de confiance concordent avec ceux de SAS et STATA mais pas avec ceux d'Epi-Info, qui sont trop étroits.

170. Le tableau XXI.3, rang 5, de l'annexe (CD-ROM) montre que WesVar concorde avec tous les autres logiciels sur les estimations ponctuelles des domaines et les erreurs types estimatives. Les intervalles de confiance sont très proches de ceux de SAS SURVEYMEANS mais s'écartent légèrement de ceux de STATA svytab (utilise logit transform) et d'Epi-Info (utilise $z = 1,96$ au lieu de Student-t value).

171. Le tableau XXI.4, rang 5, de la section VII de l'annexe (CD-ROM), montre le résultat du contraste linéaire de WesVar pour comparer la prévalence de la séropositivité entre femmes rurales et urbaines. WesVar concorde avec SAS SURVEYREG, SUDAAN DESCRIPT et STATA svylogit sur l'erreur type estimative pour le contraste linéaire et Student-t statistic. Les intervalles de confiance sur le contraste linéaire présentent des différences négligeables entre SAS, STATA et WesVar.

172. Le tableau XXI.5, rangs 6 et 7, de la section VII de l'annexe (CD-ROM), montre les deux tests de chi-carré de Rao-Scott pour les données d'enquêtes par sondage complexes tels qu'ils sont effectués avec WesVar. Ces calculs ne concordent exactement avec aucun des tests de chi-carré des autres logiciels.

173. Le tableau XXI.6, rang 4 de la section VII de l'annexe (CD-ROM), montre que la procédure de régression logistique de WesVar produit le même odds-ratio estimatif et essentiellement le même intervalle de confiance que SUDAAN CROSSTAB et STATA svylogit. Le tableau XXI.6, rang 5, montre que le ratio de prévalence estimé par WesVar (utilisation des fonctions cellules de TABLES) concorde avec ceux de SUDAAN CROSSTAB et Epi-Info, avec des différences négligeables entre les intervalles de confiance de SUDAAN et de WesVar.

4. Avantages/inconvénients/coût

174. WesVar utilise uniquement les techniques de réplification pour l'estimation de la variance. Les analystes des ensembles de données secondaires à publier avec pondérations des répliqués n'ont pas besoin de connaître les détails de conception de l'échantillon (par exemple, les variables de conception de l'enquête *STRATVAR* et *PSUVAR*), mais ils doivent indiquer à WesVar la méthode utilisée pour obtenir des pondérations des répliqués (ces informations figurent dans la documentation de l'enquête). Si l'utilisateur a besoin d'utiliser WesVar pour calculer les pondérations des répliqués pour l'ensemble de données d'une enquête par sondage, il a besoin d'avoir une certaine connaissance des méthodes de réplification et il doit également connaître les trois variables de conception de l'enquête liées au plan d'échantillonnage courant *WR*

(variable de stratification *STRATVAR*, variable PSU *PSUVAR* de chaque strate, variable de pondération de l'échantillon *WTVAR*).

175. WesVar offre une vaste capacité de calcul des pondérations de répliqués pour un ensemble de données d'enquête par sondage. Il dispose de cinq différentes techniques de réplification, y compris de la possibilité d'ajustement pour les cas de non-réponse et de post-stratification ou de ratissage. En outre, WesVar permet d'incorporer un terme de correction de la population finie pour l'échantillonnage à une phase utilisant les techniques « jackknife » pour l'estimation de la variance.

176. Pour ceux qui ne maîtrisent pas encore les techniques de réplification pour l'estimation de la variance, l'appendice A du *Guide de l'utilisateur* de WesVar offre un excellent aperçu de la théorie et de la pratique de ces techniques, quoique la lecture de ce matériel nécessite une certaine connaissance des statistiques mathématiques. Par ailleurs, l'appendice D du *Guide de l'utilisateur* fournit des conseils très utiles et plusieurs exemples pour choisir une méthode de réplification pour un plan d'échantillonnage donné.

177. WesVar permet d'estimer certaines fonctions des paramètres de population définies par l'utilisateur, ce qui est plus difficile à faire avec l'approche de l'estimation de la variance par la linéarisation par série de Taylor. Il est donc intrinsèquement plus souple que les autres logiciels examinés dans ce chapitre pour ce qui est de sa capacité d'estimation des paramètres de population. Bien que SUDAAN offre les méthodes de réplification BRR et « jackknife » pour l'estimation de la variance, à la différence de WesVar, SUDAAN ne permet pas à l'utilisateur de spécifier les fonctions des paramètres de population à estimer.

178. Il est parfois plus difficile d'utiliser les résultats directs de WesVar que ceux de la plupart des autres logiciels d'enquête par sondage. Ces résultats comprennent un rang pour chaque cellule d'un tableau demandé (comme le montre la section VI de l'annexe). Cependant, il offre une utilité « visualiseur de tableau » qui peut être téléchargée gratuitement du site web de WesVar. Ce complément de programme convertit les résultats de WesVar 4 en une grille ou un tableau pouvant être visualisé sur l'écran ou imprimé ou produit un fichier électronique qui peut être « collé » à des logiciels tels que Microsoft Word ou Excel.

179. Comparé à celui des autres logiciels examinés dans ce chapitre, le coût de WesVar est faible. La version 4 peut être téléchargée gratuitement pour une période d'essai de trente jours, et la version 2 pour un usage illimité.

L. PC-CARP

180. PC-CARP est un programme MS-DOS indépendant mis au point et offert par l'Université de l'État de l'Iowa (Département de statistiques). Il utilise le plan d'échantillonnage courant *WR* mentionné précédemment et, pour les conceptions plus simples, il peut incorporer les termes *fpc* jusqu'à deux phases d'échantillonnage. Il utilise la linéarisation par série de Taylor pour l'estimation de la variance.

181. Il permet de calculer les estimations ponctuelles, les erreurs types estimatives et les intervalles de confiance pour les totaux, les moyennes, les proportions, les quantiles, les fonctions de distribution empirique, les ratios et les différences de ratio de population et de catégories de population (et, par conséquent, les différences entre moyennes, proportions et totaux). Il permet également une régression linéaire fondée sur la conception et l'analyse de tableaux de contingence à deux entrées, y compris le test de chi-carré. Il permet de calculer l'effet de conception et le coefficient de variation pour les estimations ponctuelles. Il offre trois modules complémentaires : PC-CARPL

pour la régression logistique fondée sur la conception, POSTCARP pour la post-stratification de données d'enquête par sondage, et EV-CARP pour l'analyse de régression avec erreur de mesure dans les variables explicatives.

182. L'interface avec l'utilisateur se fait par navigation par clavier sur écrans de menu, sans utilisation de la souris. Il n'accepte que les fichiers ASCII comme entrée lorsque les données d'entrée risquent d'être délimitées dans leur espace ou de longueur fixe avec instruction de format en syntaxe FORTRAN. Il n'y a pas de restrictions quant au nombre d'observations de l'ensemble de données, et la plupart des analyses peuvent accepter jusqu'à 50 variables. PC-CARP peut fonctionner sur de vieux ordinateurs avec DOS 5.0 ou des ordinateurs plus récents utilisant Windows 3.1x, Windows 95 ou des versions ultérieures. Il n'occupe que 3 mégabytes (Mb) sur le disque dur et que 450 kilobytes (kb) de la mémoire vive (RAM). Tout système plus récent doit pouvoir accepter des programmes DOS pour utiliser PC-CARP.

183. Le prix d'achat à ne payer qu'une fois pour PC-CARP est faible, comparé à ceux d'autres logiciels examinés. Aucun droit de renouvellement annuel n'est exigé. Les trois modules complémentaires sont d'un prix modique.

184. Ce chapitre ne présente aucune analyse d'exemple avec PC-CARP de l'enquête effectuée au Burundi.

M. CENVAR

185. CENVAR est l'un des composants d'un logiciel statistique complet appelé Système de traitement micro-informatique intégré [*Integrated Microcomputer Processing System* (IMPS)] qui a été conçu par l'United States Bureau of the Census pour le traitement, la gestion et l'analyse de données d'enquêtes par sondage complexes. IMPS, y compris CENVAR, peut être téléchargé gratuitement du site <http://www.census.gov/ipc/www/imps/download.htm>. Au début de 2003, IMPS était utilisable en partie sur Windows et en partie encore sur DOS. Ce chapitre ne présente pas d'analyse d'IMPS.

186. CENVAR est adapté de PC-CARP et présente donc un grand nombre des caractéristiques de celui-ci. CENVAR est utilisable avec les mêmes échantillons que PC-CARP, c'est-à-dire le plan d'échantillonnage courant *WR* ainsi que l'incorporation des termes *fpc* dans l'estimation de la variance pour les échantillons plus simples à une ou deux phases avec utilisation d'échantillonnage de remplacement. Il utilise la linéarisation par série de Taylor pour l'estimation de la variance. Ce logiciel est commandé par menu et n'utilise pas la souris.

187. Il permet de calculer les estimations ponctuelles, les erreurs types estimatives, les intervalles de confiance, les coefficients de variation et les effets de conception pour les totaux, les moyennes, les proportions, les ratios et les différences de ratio (et par conséquent les différences de moyenne, de proportion et de total) de population et de catégories de population. Les autres options de PC-CARP, à savoir la régression linéaire fondée sur la conception, une analyse de tableau de contingence à deux entrées et l'estimation par quantiles, ne sont pas incluses, pas plus que les modules complémentaires de PC-CARP.

188. Le *Guide de l'utilisateur* de CENVAR (1995), d'une centaine de pages, peut être téléchargé de l'Internet. Il offre des exemples utiles et des exercices d'application tirés de trois enquêtes par sondage effectuées par le Bureau of the Census. CENVAR n'accepte que les données ASCII et nécessite le logiciel de Dictionnaire des données IMPS. Ce dictionnaire doit être créé avant l'utilisation de CENVAR. Autrement dit, il faut avoir une certaine connaissance de l'IMPS pour pouvoir utiliser CENVAR.

CENVAR fonctionne sur DOS 3.2 ou un système d'exploitation plus récent. Il prend 10 mégabytes (Mb) de mémoire sur le disque dur et 640 kilobytes (kb) de mémoire vive. Ce chapitre ne présente aucune analyse d'exemple avec CENVAR de l'enquête effectuée au Burundi.

N. IVEware (Version bêta)

189. IVEware (*Imputation and Variance Estimation Software*) [logiciel d'imputation et d'estimation de variance] est une application appellable de SAS pour les données d'enquêtes par sondage conçue récemment par le Survey Methodology Program de l'Université du Michigan. Il utilise le plan d'échantillonnage courant *WR* et la linéarisation par série de Taylor ou les méthodes de réplcation, selon la procédure.

190. Le module IMPUTE utilise l'approche de régression séquentielle multivariée pour imputer des valeurs manquantes, y compris des ensembles multiples de données imputées. Le module DESCRIBE estime les moyennes et les proportions de population et de catégories de population, les différences entre sous-groupes et les contrastes linéaires des moyennes et des proportions. La linéarisation par série de Taylor est utilisée. Le module REGRESS convient pour plusieurs modèles de régression fondés sur la conception (linéaire, logistique, etc.); la technique de réplcation « jackknife » est utilisée. Le module SASMOD permet aux utilisateurs de tenir compte des caractéristiques de conception d'échantillons complexes lorsque l'on utilise plusieurs SAS PROCs pour l'analyse des données, par exemple CATMOD, GENMOD et MIXED. Une analyse d'imputations multiples peut être effectuée par les trois modules d'analyse de données (DESCRIBE, REGRESS, SASMOD).

191. IVEware fonctionne avec SAS V6.12 ou plus élevé et peut être utilisé sur des ordinateurs individuels utilisant les systèmes d'exploitation Microsoft Windows ou Linux; il peut également être utilisé avec d'autres systèmes. Bien que les utilisateurs n'aient pas besoin de connaître les éléments constitutifs du langage SAS Macro, C et FORTRAN, il leur faut au moins un niveau modeste d'expérience de SAS et, bien entendu, du logiciel SAS. Le logiciel et la documentation IVEware peuvent être téléchargés gratuitement de <http://www.isr.umich.edu/src/smp/ive/>. Ce chapitre ne présente aucune analyse d'exemple avec IVEware de l'enquête effectuée au Burundi.

O. Conclusions et recommandations

192. Certains analystes de données peuvent être surpris de constater qu'il faut un logiciel spécialisé pour l'estimation de la variance sur les données d'enquêtes par sondage complexes. Certains voudront peut-être utiliser un logiciel conçu pour les échantillons aléatoires simples pour estimer la variance de données d'enquêtes par sondage complexes, mais nous le leur déconseillons. Il existe à présent plusieurs options de logiciel pour l'estimation de la variance, dont certains sont gratuits. Les raisons du choix entre ces options sont généralement la connaissance du logiciel, le coût, la facilité d'utilisation et la question de savoir si l'on ne cherche que des analyses descriptives de base ou si l'on veut pouvoir faire des analyses plus complètes.

193. Si vous utilisez déjà un logiciel général de statistiques offrant la possibilité d'estimation de la variance dans une enquête par sondage, dès lors, le choix est évident, car le coût d'acquisition est déjà payé et la syntaxe connue. Les utilisateurs de STATA dispose d'une vaste capacité d'estimation de variance d'enquêtes par sondage et ne devraient pas avoir à chercher plus loin, à moins que l'ensemble de données analysé ne doivent utiliser les méthodes de réplcation. Les utilisateurs de SAS, avec la Version 9.0 lancée récemment, disposent d'une capacité accrue d'estimation de la va-

riance dans les enquêtes par sondage par rapport à la Version 8.2 et peuvent compter sur un nouvel accroissement de capacité à l'avenir. Toutefois, si SAS V9.0 ne suffit pas pour vos besoins d'estimation de la variance dans les enquêtes par sondage, l'utilisation du logiciel gratuit IVEware avec SAS peut répondre à vos besoins. Les utilisateurs d'Epi-Info ne disposent que d'une capacité de base d'estimation de la variance des données d'enquêtes par sondage, mais si c'est tout ce dont ils ont besoin, ils peuvent s'en contenter. Vers la fin de 2003, le logiciel d'analyse statistique SPSS amplement utilisé a lancé un module complémentaire pour enquêtes par sondage complexes qui en fait à présent un choix viable.

194. Si votre logiciel général de statistiques ne présente pas la capacité nécessaire d'estimation de la variance des données d'enquêtes par sondage, vous pouvez envisager un logiciel d'enquête par sondage spécialisé (tel que WesVar, SUDAAN, PC-CARP ou CENVAR) ou un logiciel général de statistique différent (tel que STATA ou SAS avec/sans IVEware ou SPSS ou peut-être Epi-Info). SUDAAN attire souvent les utilisateurs de SAS parce qu'il a une syntaxe qui ressemble à celle de SAS et il peut utiliser SUDAAN appelable dans SAS, quoique lorsqu'il est utilisé en mode indépendant, il accepte également les ensembles de données d'entrée SPSS. WesVar, PC-CARP et CENVAR sont tous trois des programmes indépendants ayant leur propre organisation unique, de sorte que la connaissance de certains autres logiciels de statistiques a peu de chances d'influer sur le choix entre ces trois programmes. PC-CARP et CENVAR peuvent plaire à ceux qui doivent ou préfèrent utiliser DOS, et ne pas plaire à ceux qui préfèrent Windows.

195. Si le coût est un important facteur dans le choix du logiciel, certains programmes sont nettement préférables à d'autres. Epi-Info, quoique gratuit, est limité dans les options analytiques pour l'estimation de la variance, mais tout à fait acceptable pour les analyses de base. CENVAR, gratuit lui aussi, offre davantage d'options analytiques qu'Epi-Info mais pas de procédures de régression fondée sur la conception. WesVar Version 2 est également gratuit. IVEware est gratuit, mais il doit être utilisé avec SAS. Parmi les logiciels d'enquête par sondage peu coûteux mais complets figure WesVar Version 4 et PC-CARP. STATA et SUDAAN indépendant sont d'un coût modéré, et SAS est coûteux.

196. La méthode utilisée pour l'estimation de la variance est un autre facteur qui peut influencer sur le choix du logiciel. Par exemple, si vous analysez un ensemble de données destiné à être publié comprenant des pondérations de répliqués BRR ou « jack-knife » et pas de variables d'identification des strates/UPE, un logiciel qui n'utilise que la linéarisation par série de Taylor ne vous sera d'aucune utilité. Parmi les logiciels examinés ici, SUDAAN et IVEware offrent tous deux la linéarisation par série de Taylor et les méthodes de réplification. WesVar n'offre que les procédures de réplification, et STATA, SAS, PC-CARP, Epi-Info et CENVAR offrent la linéarisation par série de Taylor.

197. Enfin, le choix du logiciel dépend des analyses que vous désirez effectuer. Tous les huit logiciels examinés ici peuvent effectuer des analyses de base et des analyses descriptives. Parmi ceux de ces huit logiciels qui vont au-delà des analyses de base figurent STATA, SUDAAN, WesVar, PC-CARP et SAS (avec ou sans IVEware). Le tableau XXI.2 résume et compare de nombreuses caractéristiques de ces huit logiciels.

198. Les cinq logiciels comparés de façon empirique dans ce chapitre (SAS, SUDAAN, STATA, Epi-Info et WesVar) permettent, comme prévu, les mêmes estimations ponctuelles pour tous les exemples descriptifs et analytiques. Ces cinq logiciels produisent essentiellement les mêmes erreurs types estimatives, que ce soit avec BRR ou avec la linéarisation par série de Taylor. Il y a de légères variations entre ces logiciels

sur certains des calculs de l'intervalle de confiance, dont les raisons ont été examinées précédemment. Autrement dit, il n'y a pas de raison impérieuse de choisir l'un de ces logiciels plutôt qu'un autre sur la base des analyses comparatives présentées dans ce chapitre.

199. Le marché des logiciels spécialisés d'enquête par sondage (axés sur l'estimation de la variance) pourrait disparaître à l'avenir. La tendance semble être à inclure ces capacités dans les logiciels standard (par exemple, STATA, SAS et SPSS). Ainsi, à l'avenir, il pourrait être plus aisé aux analystes de données d'obtenir et d'utiliser un logiciel approprié pour l'estimation de la variance de données d'enquêtes complexes.

REMERCIEMENTS

Nos remerciements s'adressent à :

Michael S. Deming, MD, MPH, pour nous avoir fourni l'ensemble de données sur le Burundi et sa documentation, pour sa lecture attentive des multiples versions de notre manuscrit et pour ses précieuses suggestions d'édition.

Kevin Sullivan, PhD, pour ses conseils et ses précieuses suggestions à propos de l'utilisation d'Epi-Info, pour sa lecture attentive des multiples versions de notre manuscrit et pour ses précieuses suggestions d'édition.

Z. T. Daniels, MS, MBA, pour le formatage des résultats et des tableaux à texte de WesVar et pour avoir placé les données sur la population du Burundi sur l'Internet.

Graham Kalton, PhD, et Ibrahim Yansaneh, PhD, pour leurs précieux conseils d'organisation et leurs suggestions d'édition.

James Chromy, PhD, pour sa lecture attentive des diverses versions du manuscrit et ses précieuses suggestions d'édition.

Plusieurs conseillers anonymes pour leur lecture attentive des diverses versions du manuscrit et leurs précieuses suggestions d'édition.

Paul Weiss, MS, pour ses instructions et ses précieuses suggestions concernant l'utilisation de WesVar.

Toutes erreurs contenues dans ce chapitre n'engagent que la responsabilité de l'auteur.

ANNEXE

Voir note en début de chapitre, p. 421.

RÉFÉRENCES

Brogan, Donna (1998 et sous presse). Software for sample survey data: misuse of standard packages. Invited chapter in *Encyclopedia of Biostatistics*, Peter Armitage et Theodore Colton, rédacteurs en chef. New York, John Wiley, vol. 5, pp. 4167-4174. Chapitre révisé sous presse pour la 2^e édition de l'*Encyclopedia of Biostatistics*, dont la publication est prévue pour 2004.

_____, *et al.* (1994). Increasing the accuracy of the expanded programme on immunization's cluster survey design. *Annals of Epidemiology*, vol. 4, n° 4, pp. 302-311.

Carlson, Barbara L. (1998). Software for sample survey data. Dans *Encyclopedia of Biostatistics*, vol. 5, Peter Armitage et Theodore Colton, rédacteurs en chef, New York, John Wiley and Sons, pp. 4160-4167.

- Cochran, William G. *Sampling Techniques*, 3^e éd. New York, John Wiley and Sons.
- Expanded Programme on Immunization (EPI) [1996]. Estimating tetanus protection of women by serosurvey. *Weekly Epidemiological Record*. Organisation mondiale de la santé, vol. 71, pp. 17-124.
- Hansen, Morris H., William N. Hurwitz et William G. Madow (1953). *Sample Survey Methods and Theory*, vol. I, *Methods and Applications*. New York, John Wiley and Sons.
- Judkins, D. (1990). Fay's method for variance estimation. *Journal of Official Statistics*, vol. 6, pp. 223-240.
- Kish, Leslie (1965). *Survey Sampling*. New York, John Wiley and Sons.
- _____, et M. R. Frankel (1974). Inference from complex samples. *Journal of the Royal Statistical Society, Series B*, vol. 36, pp. 1-37.
- Korn, Edward L., et Barry I. Graubard (1999). *Analysis of Health Surveys*. New York, John Wiley and Sons.
- Krotki, Karol P. (1998). Sampling in developing countries. Dans *Encyclopedia of Biostatistics*, vol. 5, Peter Armitage et Theodore Colton, rédacteurs en chef. New York, John Wiley and Sons, pp. 3939-3944.
- Levy, Paul S., et Stanley Lemeshow (1999). *Sampling of Populations: Methods and Applications*, 3^e éd. New York, John Wiley and Sons.
- Lohr, Sharon L. (1999). *Sampling: Design and Analysis*. Pacific Grove, Californie, Duxbury Press, Brooks/Cole Publishing.
- Rao, J. N. K., et A. J. Scott (1981). The analysis of categorical data from complex sample surveys: chi-squared tests for goodness of fit and independence in two-way tables. *Journal of the American Statistical Association*, vol. 76, pp. 221-230.
- _____. (1984). On chi-squared tests for multiway contingency tables with cell proportions estimated from survey data. *Annals of Statistics*, vol. 12, pp. 46-60.
- Rust, K. F., et J. N. K. Rao (1996). Variance estimation for complex surveys using replication techniques. *Statistical Methods in Medical Research*, vol. 5, pp. 283-310.
- Shah, Babubhai V. (1998). Linearization methods of variance estimation. In *Encyclopedia of Biostatistics*, vol. 3, Peter Armitage et Theodore Colton, rédacteurs en chef. New York, John Wiley and Sons, pp. 2276-2279.
- Som, R. K. (1995). *Practical Sampling Techniques*, 2^e éd. New York, Basel et Hong Kong, Marcel Dekker.
- Wolter, Kirk M. (1985). *Introduction to Variance Estimation*. New York, Springer-Verlag.

Références pour les logiciels :

- CENVAR Variance Calculation System: IMPS Version 3.1: *User's Guide*, 1995. Bureau of the Census, United States Department of Commerce, Washington, D.C. Peut être consulté sur <http://www.census.gov/ipc/www/imps/download.htm>.
- Epi-Info. Peut être consulté sur <http://www.cdc.gov/epiinfo/> pour le logiciel et la documentation.
- IVEware. Peut être consulté sur <http://www.isr.umich.edu/src/smp/ive/> pour le logiciel et la documentation.
- PC CARP (1986, 1989). *User's Manual*, Wayne Fuller and others, eds. Statistical Laboratory, Iowa State University, Ames, Iowa. Peut être consulté sur <http://cssm.iastate.edu/software>.
- Research Triangle Institute (2001). *SUDAAN User's Manual, Release 8.0*. Research Triangle Park, Caroline du Nord, Research Triangle Institute. Peut être consulté sur www.rti.org/sudaan.

SAS/STAT. Peut être consulté sur <http://www.sas.com/technologies/analytics/statistics/stat/index.html> pour des informations sur le logiciel SAS/STAT qui comprend des procédures pour données d'enquêtes par sondage.

STATA. Peut être consulté sur <http://www.stata.com> pour STATA, sur <http://www.stata.com/help.cgi?svy> pour un examen des commandes svy dans STATA, et sur <http://www.stata.com/bookstore/> pour la disponibilité de manuels de référence.

WesVar 4.2-*User's Guide* (2002). Rockville, Maryland, Westat. Voir aussi le site web <http://www.westat.com/WesVar/about/>.

Partie 2

Études de cas

Introduction

GAD NATHAN

Université hébraïque
Jérusalem, Israël

1. Dans la première partie de la présente publication, nous nous sommes efforcés de présenter les derniers progrès réalisés sur les aspects les plus importants de la conception et de l'analyse des enquêtes sur les ménages dans les pays en développement et les pays en transition. Nous avons concentré notre attention sur les principes généraux et les méthodologies de conception, mise en œuvre et analyse des enquêtes sur les ménages dans les pays en développement et les pays en transition, en mettant l'accent sur les caractéristiques opérationnelles : effets de conception, coûts des enquêtes et erreurs autres que les erreurs d'échantillonnage. Une grande variété de méthodes et de techniques ont été conçues et appliquées aux enquêtes sur les ménages dans les pays en développement et les pays en transition. Les domaines couverts dans les précédents chapitres ont donc été aussi vastes que possibles, de manière à assurer le traitement du plus grand nombre de cas possible. De nombreux exemples ont été inclus dans ces chapitres eux-mêmes, et certaines applications spécifiques à diverses enquêtes effectuées dans des pays en développement et des pays en transition ont été traitées dans des chapitres particuliers. C'est ainsi que le chapitre VII a décrit les conceptions d'échantillons et présenté des données sur les effets de conception observés dans 11 enquêtes effectuées dans sept pays. De même, le chapitre XI a présenté de façon détaillée une étude de cas sur les pratiques actuelles concernant la notification, le contrôle, l'évaluation et la compensation des erreurs autres que les erreurs d'échantillonnage au Brésil.

2. Toutefois, pour les praticiens, il est essentiel de voir comment les diverses techniques et méthodes préconisées peuvent être combinées dans la pratique en une application dans la réalité et d'examiner des exemples concrets de l'intégration des méthodes en une enquête complète sur les ménages bien conçue et bien analysée. Les conditions propres à chaque pays et son infrastructure ont une profonde influence sur la façon dont les principes généraux s'appliquent dans la pratique et, en particulier, sur la façon dont ils sont combinés pour donner lieu à une enquête approfondie. En règle générale, les études de cas sont un outil d'apprentissage indispensable à l'étude de toute science appliquée, et l'étude de l'application de concepts statistiques théoriques et de leurs résultats à la conception et à l'analyse d'enquêtes statistiques au moyen d'études détaillées est particulièrement utile. C'est pour cette raison que nous avons consacré la seconde partie de cette publication à des études de cas. Avec ces études, nous espérons présenter les méthodes examinées dans la première partie dans leur application à la réalité. Il ne s'agit pas d'illustrer simplement l'application de certains aspects donnés des techniques étudiées mais, surtout, leur intégration dans des programmes complexes de conception et d'analyse des enquêtes sur les ménages dans les pays en développement et les pays en transition.

3. Les quatre chapitres de cette partie de notre publication couvrent un large éventail de plusieurs centaines d'enquêtes sur les ménages effectuées dans le monde entier sur divers thèmes, dans différentes conditions et selon différentes conceptions, et à divers degrés de détail. Dans la plupart des cas, ces études décrivent les buts et la portée de ces enquêtes, la définition de la population et la conception de l'échantillon, les instruments utilisés, la conception et l'exécution du travail de terrain, les erreurs dues aux cas de non-réponse et l'évaluation, l'analyse, la pondération et les effets de conception. Dans certains cas, les paramètres de conception des enquêtes décrites ont été normalisés sur un grand nombre d'enquêtes effectuées par des organisations internationales. Dans d'autres cas, les enquêtes conçues présentaient des similarités dues à la similarité des conditions entre pays voisins (notamment pour les pays en transition).

4. Le chapitre XXII décrit les caractéristiques générales et la conception du Programme d'enquêtes démographiques et sanitaires [*Demographic and Health Surveys* (DHS) programme] qui a comporté plus de 100 enquêtes sur les ménages et auprès de particuliers effectuées dans plus de 50 pays. Le chapitre XXIII décrit les caractéristiques opérationnelles de la série de plus de 60 enquêtes effectuées dans plus de 40 pays sous l'égide de la Banque mondiale pour l'Étude sur la mesure des niveaux de vie (LSMS). Le chapitre XXIV examine un certain nombre d'échantillons et de questions liées à la mesure dans les Enquêtes nationales sur le budget des ménages (HBS), sur la base de l'expérience de ces enquêtes dans un certain nombre de pays en développement et de pays en transition. Une étude de cas consacrée à l'Enquête de 1997-1998 sur les dépenses et la consommation en République démocratique populaire lao donne des descriptions détaillées des conditions générales d'exécution du travail d'enquête, des instruments de l'enquête, des méthodes de mesure, de la conception de l'échantillon et du travail de terrain. Enfin, le chapitre XXV examine les principaux aspects de la conception et de l'exécution d'enquêtes sur les ménages dans 14 pays en transition d'Europe orientale, avec des descriptions détaillées des études de cas des enquêtes sur les ménages effectuées dans six de ces pays.

5. Certaines des caractéristiques décrites ont beaucoup en commun. Par exemple, toutes les enquêtes étaient des enquêtes sur les ménages ou avaient en elles un élément de ménage. Toutefois, dans beaucoup de cas, l'unité analysée était essentiellement l'individu, un membre d'une famille (par exemple, les femmes dans les enquêtes démographiques et sanitaires) ou tous les membres de la famille (par exemple, les enquêtes sur la main-d'œuvre), et souvent, la réponse était donnée par un intermédiaire. Les échantillons de base étaient très semblables dans la plupart des enquêtes décrites, échantillonnage en grappe à phases multiples généralement avec de vastes unités géographiques comme unités primaires d'échantillonnage (UPE). Parfois, un essai de stratification des UPE a été tenté. Le plus souvent, les échantillons étaient autopondérés au niveau du ménage. Cependant, lorsque la sélection portait sur un individu par ménage, l'échantillon d'individus n'était plus autopondéré. Pratiquement tous les échantillons étaient de conception pleinement probabiliste, quoique les enquêtes sur le budget des ménages effectuées en République tchèque et en Slovaquie continuent d'utiliser l'échantillonnage par quotas.

6. Les enquêtes ont des buts et des propos très variés. Par exemple, les Enquêtes démographiques et sanitaires visent à « fournir aux pays les données nécessaires pour suivre et évaluer leurs programmes de population, de santé et de nutrition ». Le programme LSMS a pour objet de chercher à comprendre, mesurer et suivre les conditions de vie. Le Programme d'enquêtes sur le budget des ménages tend à mesurer les aspects importants du budget quotidien, revenu et dépenses, des ménages. Les enquêtes très variées effectuées sur les ménages des pays en transition cherchent plus

particulièrement à analyser les conditions de vie, la construction de l'indice des prix à la consommation et les statistiques de main-d'œuvre nécessaires pour le passage d'une économie étatisée à l'économie de marché.

7. Les instruments utilisés dans ces enquêtes restent fondés, d'une façon générale, sur des entrevues sur le terrain, avec questionnaires sur papier et crayon. Toutefois, une première tentative d'utilisation d'entrevues téléphoniques assistées par ordinateur [computer- assisted telephone interviewing (CATI)] a été signalée à propos de l'enquête sur la main-d'œuvre estonienne (chapitre XXV). Un grand nombre des enquêtes examinées ont accordé une grande priorité à la formation et au contrôle des enquêteurs, afin de réduire le nombre de cas de non-réponse et le nombre d'erreurs de réponse. Des taux de réponse élevés ont été signalés pour les DHS : 87-99% pour les femmes. Les LSMS ont également affiché des taux globaux de réponse élevés (74-99,7%), mais aussi des taux élevés de données manquantes sur le revenu, surtout de la part des travailleurs indépendants. L'Enquête sur le budget des ménages de la République démocratique populaire lao n'a affiché que 3,1% de non-réponse. En revanche, les enquêtes sur le budget des ménages effectuées dans les pays en transition ont signalé des taux de non-réponse de 4 à 29%, et certains pays ont régulièrement affiché des taux de non-réponse de moins de 10%.

8. Beaucoup d'études de cas mettent l'accent sur les efforts consacrés à nettoyer, éditer et imputer les données. La majeure partie du traitement et de l'analyse ont été effectués à l'aide de logiciels courants – souvent sans pondération. Les pays en transition, pour leur part, ont abondamment utilisé la pondération et l'étalonnage. Un grand nombre des études ont cherché à estimer les effets de conception en utilisant des méthodes courantes. Ces estimations ont servi à la fois pour l'analyse et pour améliorer la conception pour les études à venir. C'est ainsi qu'un examen des effets de conception des LSMS a révélé la nécessité d'utiliser ces études pour l'analyse, mais les grandes différences observées sur différentes variables importantes dans les effets de conception n'ont pas permis de parvenir à des conclusions utiles sur la conception de l'échantillon, en raison de la multiplicité des sujets des enquêtes.

9. Par-delà les possibilités d'apprendre que permet de tirer le large éventail d'expériences présentées ici pour des enquêtes très diverses effectuées dans des pays différents, les rapports offrent leurs propres conclusions importantes sur les divers types d'enquêtes couverts. Au nombre de ces conclusions figurent la nécessité d'actualiser constamment les cadres d'échantillonnage, de mettre constamment l'accent sur la formation sur le terrain et le contrôle des enquêteurs, l'importance d'une bonne préparation des données, la formulation et l'actualisation des besoins et de l'analyse des données, l'utilisation des effets de conception, et plus encore. Considérées conjointement avec les méthodes décrites dans la première partie de cette publication, ces études de cas constituent un élément important et inhérent des enseignements que l'on peut tirer de la présente publication.

Chapitre XXII

Les Enquêtes démographiques et sanitaires

MARTIN VAESSEN
 ORC Macro
 Calverton, Maryland
 États-Unis d'Amérique

MAMADOU THIAM*
 UNESCO
 Montréal, Canada

THANH LÊ*
 WESTAT
 Rockville, Maryland
 États-Unis d'Amérique

* Mamadou Thiam et Thanh Lê travaillaient auparavant au programme DHS à ORC Macro.

RÉSUMÉ

Le présent chapitre offre un aperçu des principales procédures suivies dans le cadre du Programme international d'enquêtes démographiques et sanitaires [*International Demographic and Health Surveys (DHS) programme*] pour l'exécution de grandes enquêtes sur les ménages et les personnes. Il donne une description d'ensemble du contenu général de ces enquêtes, des procédures d'échantillonnage, des taux de réponse et des effets de conception, ainsi que des procédures et approches suivies pour toutes les composantes importantes de ces enquêtes, depuis la formation jusqu'au traitement des données et à la rédaction de rapports. Il présente également une récapitulation des principales leçons tirées jusqu'ici de l'exécution de ce programme d'enquêtes.

Mots clés : enquêtes sur les ménages, taux de réponse, échantillonnage, erreurs d'échantillonnage, effets de conception, travail de terrain.

A. Introduction

1. Le Programme d'enquêtes démographiques et sanitaires (DHS) effectue des enquêtes sur les ménages dans les pays en développement depuis 1984. Ces enquêtes ont principalement pour but de doter les pays des données nécessaires pour suivre et évaluer régulièrement leurs programmes de population, de santé et de nutrition. L'absence de données appropriées à tirer des statistiques administratives et d'autres systèmes courants de collecte de données explique le fait que les donateurs et les pays attachent une importance croissante à l'utilisation d'indicateurs objectifs pour mesurer les progrès réalisés et le fait qu'ils utilisent de plus en plus les données des enquêtes sur les ménages. Dans une enquête démographique et sanitaire, un échantillon de ménages est sélectionné à travers tout le pays, puis ces ménages sont interrogés au moyen d'un questionnaire destiné à recueillir des données sur les principaux aspects de leurs logements et à identifier tous les membres de la famille et leurs principales caractéristiques. Les femmes âgées de 15 à 49 ans sont également interrogées à l'aide d'un questionnaire destiné à recueillir des informations principalement sur leurs antécédents, leur comportement procréateur, leur connaissance et leur utilisation de contraceptifs, leur santé et celle de leurs enfants et d'autres questions. La durée moyenne des entrevues est d'environ 35 à 40 minutes, mais elles peuvent durer entre 10 et 90 minutes, quoique certaines durent plus longtemps. La taille des échantillons est très variable et

va de 5 000 à 30 000 femmes. Dans certains pays, un échantillon d'hommes âgés de 15 à 49 ans est également interrogé. Souvent, il s'agit d'un sous-échantillon de l'échantillon utilisé pour sélectionner les femmes. Les entrevues avec les hommes durent en moyenne 25 minutes. Les sections suivantes présentent l'historique du Programme DHS ainsi que le contenu général des enquêtes, un aperçu de ses procédures d'échantillonnage et une analyse des cas de non-réponse unitaire. Les effets de conception de l'échantillonnage sont également présentés, ainsi que les différentes phases de l'exécution de l'enquête et les leçons tirées des Enquêtes démographiques et sanitaires dans les pays en développement.

B. Historique

2. Les Enquêtes démographiques et sanitaires font suite à deux autres programmes d'enquêtes sur les ménages : les Enquêtes mondiales sur la fécondité [*World Fertility Surveys* (WFS)] et les Enquêtes sur la prévalence de la contraception [*Contraceptive Prevalence Surveys* (CPS)]. Les Enquêtes mondiales sur la fécondité se sont déroulées de 1973 à 1984, et les Enquêtes sur la prévalence de la contraception de 1977 à 1985. Le programme WFS a permis d'effectuer des enquêtes dans 41 pays en développement et de collaborer à des enquêtes dans 20 pays développés. Les Enquêtes mondiales sur la fécondité visaient principalement à recueillir des informations sur la fécondité, la planification familiale et, dans une certaine mesure, la santé de l'enfant. Ce programme était financé conjointement par l'United States Agency for International Development (USAID) et par le Fonds des Nations Unies pour la population (FNUAP), avec l'appui des Gouvernements du Royaume-Uni de Grande-Bretagne et d'Irlande du Nord, des Pays-Bas et du Japon.

3. Le Programme CPS a porté sur l'exécution de 43 enquêtes dans 33 pays et était axé plus spécifiquement sur la planification familiale. Il était financé par l'USAID, et les enquêtes étaient limitées aux pays qui avaient reçu une aide au développement de l'USAID.

4. Les Enquêtes démographiques et sanitaires ont commencé en 1984. À la fin de 2003, 150 enquêtes avaient été effectuées sur les femmes, 75 sur les hommes et 10 sur les établissements de santé dans quelque 70 pays. Ces enquêtes ont lieu généralement tous les cinq ans, quoique, dans certains pays, elles soient plus rapprochées. Elles se déroulent principalement dans les pays qui reçoivent une aide de l'USAID, mais certains pays ont pu participer au programme grâce à un financement de la Banque mondiale ou du FNUAP. Dans beaucoup de pays, ces enquêtes bénéficient de l'appui d'autres donateurs que l'USAID, tels que le Department for International Development (DFID) du Royaume-Uni, le Fonds des Nations Unies pour l'enfance (UNICEF) et les Gouvernements du Japon et de la Suède, entre autres. Les Enquêtes démographiques et sanitaires donnent un panorama complet des questions de population et de santé maternelle et infantile dans les pays participants, et les données sont librement accessibles aux fins de suivi et d'évaluation. Le contenu de ces enquêtes a changé au fil des années, s'adaptant à l'évolution des circonstances et des priorités.

C. Contenu

5. Le noyau de base de chaque enquête démographique et sanitaire est le même, de manière à maximiser la comparabilité des informations. En plus de ce noyau de base, les pays peuvent choisir d'ajouter des modules traitant de questions qui les intéressent plus particulièrement. Le noyau de base des questionnaires pour les pays

d'Afrique subsaharienne est quelque peu différent de ceux utilisés pour les autres pays, essentiellement du point de vue de sa complexité.

6. Le noyau de base des questionnaires utilisés durant les années 1997-2002 portaient sur les éléments suivants :

Questionnaire sur les ménages. Ce questionnaire visait à recueillir des données de base sur l'âge, le sexe, les parents vivants et la scolarité des membres du ménage. Il comprenait également des questions sur l'approvisionnement en eau et les éléments de confort du ménage, ainsi que sur la taille et le poids des femmes âgées de 15 à 49 ans et des enfants de moins de cinq ans et leurs niveaux d'hémoglobine pour la mesure de l'anémie.

Questionnaire sur les femmes. Ce questionnaire, appliqué aux femmes en âge de procréer, comprenait les sections suivantes :

- Antécédents de la répondante;
- Procréation;
- Contraception;
- Grossesses, soins post-natals et allaitement au sein;
- Vaccination, santé et nutrition;
- Mariage et activité sexuelle;
- Préférences en matière de fécondité;
- Questions d'ordre général sur le mari et le travail de la femme;
- VIH/sida et autres maladies sexuellement transmissibles.

Certaines enquêtes comprenaient un test de dépistage de la séropositivité ou de la syphilis et d'autres marqueurs biologiques.

7. Il y a aussi un questionnaire pour les hommes, qui couvre certains des mêmes thèmes que le questionnaire pour les femmes, Il n'est pas utilisé dans tous les pays. Il existe aussi un questionnaire pour prestataires de services de planification familiale et de santé, mais il ne s'inscrit pas dans le cadre de l'enquête sur les ménages et est administré aux prestataires de services. Ce questionnaire est dénommé Questionnaire d'évaluation de la prestation de services [*Service Provision Assessment (SPA) questionnaire*]. Il couvre tous les aspects de la prestation de services au moyen de questions posées aux prestataires et à leurs clients et par l'observation de la prestation des services.

8. Dans le cadre du programme DHS, il a été conçu un certain nombre de modules que les pays peuvent ajouter à leur questionnaire. Il existe des modules sur :

- La mutilation génitale des femmes;
- La mortalité maternelle;
- L'utilisation de la pilule;
- La stérilisation;
- La consanguinité (mariage entre membres d'une même famille);
- L'autopsie verbale (questions détaillées sur la cause du décès);
- Le VIH/sida;
- L'éducation des enfants;
- La condition de la femme;
- La violence domestique;

- Le paludisme;
- Les dépenses de santé des ménages.

9. En raison de la longueur du noyau de base, il est généralement impossible à un pays donné d'ajouter plus de deux ou trois modules, mais cela peut varier en fonction de la longueur des modules choisis (visiter www.measuredhs.com pour les questionnaires et autres matériels).

D. Cadre d'échantillonnage

10. Bien entendu, la question de pouvoir disposer d'un cadre d'échantillonnage approprié est traitée aux premiers stades de la planification des enquêtes démographiques et sanitaires. Ces enquêtes servent à recueillir des données sur les personnes vivant dans des familles, mais on ne dispose généralement pas d'une liste à jour de ces personnes et de ces familles. Par définition, le cadre d'échantillonnage utilisé dans la plupart des enquêtes démographiques et sanitaires est une liste d'unités géographiques couvrant l'ensemble du territoire national sans se recouper. Pour les besoins du cadre, les caractéristiques essentielles de ces unités sont des limites bien définies et des cartes clairement délimitées. Chaque unité géographique a son propre code d'identification. Elle doit aussi avoir une mesure actuelle ou estimative de sa taille (population et/ou nombre de ménages). Chaque unité est en outre classée comme urbaine ou rurale, et ces caractéristiques peuvent servir pour les besoins de la stratification.

11. Dans la plupart des pays, les unités géographiques utilisées correspondent aux zones d'énumération (ZE) qui offrent un cadre pratique pour la première phase de l'échantillonnage. Dans certains pays, ces ZE peuvent avoir une nombreuse population; dans d'autres, non. Quelle que soit leur taille, ces ZE sont généralement les unités primaires d'échantillonnage (UPE). Dans certaines enquêtes, elles sont également les unités géographiques ultimes, si elles sont suffisamment petites. Si elles servent d'UPE et sont jugées de taille trop grande (en nombre de ménages ou en population), une segmentation en tant que stade intermédiaire de sélection est alors introduite dans la conception de l'échantillon.

12. Comme on l'a mentionné précédemment, le cadre, qu'il soit composé des ZE du recensement ou d'autres unités, peut avoir perdu de son actualité. Il faut alors généralement prendre certaines mesures : soit *a*) pour actualiser le cadre tout entier; soit *b*) pour l'actualiser en partie en établissant une liste actuelle des ménages à l'avant-dernier stade de la sélection.

13. Pour certaines enquêtes, on utilise un échantillon maître préexistant comme cadre d'échantillonnage s'il est établi que sa conception peut répondre aux objectifs de mesure de l'Enquête démographique et sanitaire.

E. Phases de l'échantillonnage

14. Comme pour la conception de tout échantillon, le nombre de phases de l'échantillonnage est fonction des caractéristiques du cadre d'échantillonnage et des objectifs de l'enquête. Sans être normalisée à l'échelle internationale, la conception de l'échantillon de toute Enquête démographique et sanitaire obéit aux mêmes principes généraux : simplicité, échantillonnage probabiliste (probabilité de sélection connue non nulle), mise en grappe et stratification. Dans les enquêtes démographiques et sanitaires, il faut généralement au moins deux phases de sélection, selon la taille des unités géographiques du cadre d'échantillonnage.

15. La conception de base implique la sélection d'unités géographiques durant la première phase selon une probabilité proportionnelle à la taille, celle-ci étant définie par les chiffres de population ou le nombre de ménages de chaque unité géographique. La première phase de la sélection marque le point au-delà duquel les opérations d'échantillonnage passent du bureau au terrain pour l'établissement de cartes et, s'il y a lieu, des listes de familles des unités géographiques sélectionnées. L'établissement de cartes consiste à dessiner un croquis indiquant les limites de chaque UPE sélectionnée et l'emplacement des logements à l'intérieur de l'UPE. Dans les pays qui disposent de cartes détaillées et précises des UPE, l'établissement de cartes se limite à l'actualisation de l'implantation des logements. Lorsque le cadre n'est pas jugé tout à fait à jour, des listes actualisées des ménages sont établies pour chaque UPE sélectionnée par l'établissement de la liste de tous les ménages de chaque unité d'habitation occupée, y compris des ménages qui sont absents au moment de la visite de l'équipe chargée d'établir la liste. Les listes ainsi établies servent de cadre d'échantillonnage pour la sélection systématique de ménages durant la deuxième phase.

16. Pour chaque enquête sur les ménages, la taille de la grappe (nombre de ménages/femmes à sélectionner par UPE ou par grappe) dépend de la variable considérée. Pour les variables à forte mise en grappe, qui nécessitent souvent des comparaisons entre zones géographiques (tel est le cas de la prévalence de la contraception et de ses déterminants), la taille optimale des grappes a été évaluée entre 15 et 20 femmes par grappe. D'autres variables de fécondité ont un degré de mise en grappe moins élevé, et lorsque les comparaisons recherchées ne sont pas géographiques (par exemple, lorsqu'il s'agit de comparer différents groupes d'âge ou différents niveaux d'instruction), la taille optimale des grappes peut être plus grande. Les DHS utilisent des grappes de 30 à 40 femmes pour le secteur rural. Dans les zones urbaines, l'avantage économique d'une grappe de grande taille est généralement plus faible, et les DHS utilisent des grappes de 20 à 25 femmes. Lorsqu'il existe une liste récente des ménages, ces chiffres sont réduits car le facteur favorisant de grandes grappes est l'économie dans les opérations d'établissement de listes (ORC Macro, 1996). Les DHS collectant également des données sur la santé des enfants, et ces enfants étant ceux des femmes sélectionnées, la taille de la grappe doit être suffisamment grande pour produire le nombre d'enfants requis pour l'analyse.

17. Tous les membres admissibles des familles sélectionnées sont inclus dans l'échantillon final. Bien que, dans la plupart des échantillons utilisés dans les DHS, le nombre de ménages choisis par UPE varie d'une UPE à l'autre, dans certaines enquêtes, on a utilisé un échantillon de taille fixe.

18. Souvent, les UPE sélectionnées sont de trop grande taille pour figurer directement sur une liste. La segmentation est introduite dans la conception pour réduire le volume de travail d'établissement de listes et pour maintenir la charge de travail égale d'une UPE à l'autre. Chaque grande UPE est divisée en segments dont un est retenu dans l'échantillon selon une probabilité proportionnelle à sa taille [*probability proportional to size* (PPS)].

19. La majorité des échantillons de DHS sont en grappe et stratifiés. La stratification explicite est généralement fondée sur des critères géographiques tels que la classification urbaine/rurale et est introduite seulement durant la première phase de l'échantillonnage. Les UPE sont sélectionnées de façon indépendante dans chaque strate. La stratification implicite se fait par l'emploi de la technique de sélection automatique. En général, le nombre d'UPE est élevé, variant d'environ 300 à 550 pour un échantillon de 10 000 ménages.

20. Les DHS s'efforcent de maintenir la conception de leur échantillon aussi simple que possible, afin d'en faciliter la bonne exécution. Cependant, la conception de base est modifiée de manière à répondre aux conditions particulières du pays. Parmi ces modifications figure l'utilisation du segment de conception type avec ou sans grappes compactes; les grappes compactes sont celles où chaque ménage sélectionné est géographiquement contigu d'un autre, les ménages dispersés formant des grappes non compactes. C'est là une variante de la conception de l'échantillon où un segment type de taille prédéterminée, c'est-à-dire l'unité géographique ultime spécifiée, est aussi petit qu'il est possible. À chaque UPE ou zone d'énumération i du pays, on alloue un nombre de segments s_i en divisant sa population du recensement par la taille du segment type. Les UPE sont alors sélectionnées selon une probabilité proportionnelle à leur taille (PPS), où la mesure de la taille est égale au nombre de segments s_i . À l'intérieur de chaque UPE, un segment est alors sélectionné au hasard. L'intérêt du segment type est que les segments sont de taille moyenne T , T étant la taille désirée de la grappe. On peut ainsi éviter l'établissement de liste en utilisant l'approche de la sélection de tous les segments (ORC Macro, 1996).

21. Les estimations des DHS sont présentées à la fois pour le pays dans son ensemble et pour certains domaines géographiques, tels que domaine urbain ou rural et région. Comme la taille de la population des domaines est souvent variable, l'échantillon est généralement conçu de manière à sur-représenter les petits domaines pour offrir les tailles d'échantillon adéquates nécessaires à l'analyse. Bien entendu, cela introduit une distorsion potentielle dans les estimations nationales, que l'on corrige en pondérant comme il convient les données de l'échantillon. Le principal élément des pondérations de l'échantillon est la pondération de conception fondée sur les probabilités de sélection. On tient également compte de la non-réponse aux niveaux des ménages et des personnes dans la pondération. On peut utiliser une phase finale de pondération, où l'on effectue un ajustement post-stratification chaque fois qu'un cadre géographique périmé a été utilisé pour la sélection de l'échantillon, en utilisant des projections de population émanant de sources fiables.

F. Notification des cas de non-réponse

22. Le remplacement des unités (ménages ou personnes) qui ne répondent pas n'est pas autorisé par les DHS, ce qui les différencie de nombreuses autres enquêtes. Pour obtenir le nombre voulu d'unités sélectionnées, lors de la conception de l'échantillon, on estime les taux de non-réponse enregistré à l'occasion d'enquêtes passées ou similaires et l'on utilise ce taux pour déterminer le nombre d'unités à sélectionner. Par ailleurs, on consacre de nombreux efforts durant le travail de terrain, à obtenir des taux de réponse élevés. On procède ensuite à un examen des taux de réponse obtenus lors des DHS, y compris à une comparaison des ces taux dans le temps et à travers les différentes régions.

23. Comme on l'a mentionné précédemment, les données des DHS sont recueillies à deux niveaux : les ménages et les individus. Les personnes admissibles sont pour la plupart des femmes en âge de procréer, mais dans certains pays, des hommes âgés de 15 à 49 ans sont également interrogés. Dans les Enquêtes démographiques et sanitaires, on entend par non-réponse l'impossibilité d'interroger certains ménages ou individus sélectionnés pour l'échantillon. On mesure les taux de réponse des ménages et des individus en tenant des comptes exacts de tous les ménages et individus admissibles. Le calcul des taux de réponse utilise des codes de réponse qui sont inscrits sur les questionnaires. Le questionnaire à soumettre aux ménages identifie tous les mem-

bres admissibles de chaque ménage. Seuls les membres admis à être interrogés pour l'enquête reçoivent un questionnaire individuel.

24. Au niveau des ménages, les codes de réponse sont :

- 1H Entrevue complète;
- 2H Aucun membre de la famille ou aucun répondant compétent n'est présent au domicile;
- 3H Toute la famille est absente pour une période prolongée;
- 4H Différé;
- 5H Refusé;
- 6H Logement vacant ou adresse ne correspondant pas à un logement;
- 7H Logement détruit;
- 8H Logement non trouvé;
- 9H Autres.

Le taux de réponse des ménages est alors :

$$R_H = \frac{1H}{1H + 2H + 4H + 5H + 8H}$$

25. Dans les DHS, les ménages dont les codes sont 3H, 6H, 7H et 9H sont jugés non admissibles et ne sont donc pas inclus dans le dénominateur¹. Le code 9H est généralement converti par les superviseurs en l'un des codes explicites et est donc presque toujours inexistant. Les quelques cas de ménages restant codés 9H peuvent être classés comme non admissibles. Il convient de noter qu'en raison de l'absence dans beaucoup de pays d'un bon système d'adresses, l'établissement de listes permet d'abord d'identifier les logements par le nom des ménages qui les occupent, et ces noms sont ensuite utilisés en guise d'adresses. Le fait qu'un nouveau ménage s'installe dans un logement entre l'établissement de la liste et l'entrevue ne veut pas dire qu'une unité d'échantillonnage a été remplacée, car c'est le logement qui est vraiment la base de la sélection. De même, le fait qu'un ménage quitte le logement après l'établissement de la liste et y est remplacé par un autre ménage ne constitue pas un cas de non-réponse.

¹ Comme les ménages codés 3H (« toute la famille est absente pour une période prolongée ») sont jugés irrecevables pour les DHS, la méthode de calcul du taux de réponse des ménages est comparable à la méthode RR5 établie par l'American Association for Public Opinion Research (AAPOR) 2000 standards. Cette méthode surestime légèrement le taux de réponse vrai en ce sens qu'un petit nombre des ménages codés 3H sont admissibles mais ne sont pas inclus dans le calcul.

26. Au niveau des individus, les codes de réponse sont :

- 1I Entrevue complète;
- 2I Absent;
- 3I Différé;
- 4I Refusé;
- 5I Entrevue partiellement effectuée;
- 6I En état d'incapacité;
- 7I Autres.

Le taux de réponse des individus est :

$$R_I = \frac{1I}{1I + 2I + 3I + 4I + 5I + 6I + 7I}$$

27. Les taux de réponse non pondérés des ménages et des individus sont calculés séparément pour chaque strate ou domaine étudié et présentés dans le rapport sur la DHS du pays avec les taux de réponse globaux. Le taux de réponse global est le produit des taux de réponse aux niveaux des ménages et des individus. Dans les Enquêtes démographiques et sanitaires, les taux de réponse sont similaires dans tous les

domaines. Comme l'échantillon est généralement plus ou moins autopondéré dans chaque domaine, les taux de réponse pondérés et non pondérés du pays dans son ensemble sont très proches l'un de l'autre. Il convient de noter que les codes de réponse ci-dessus ont été utilisés dans la plupart des Enquêtes démographiques et sanitaires, mais qu'ils sont modifiés dans certaines enquêtes pour tenir compte de la situation du pays en question.

G. Comparaison des taux de non-réponse

28. À l'aide des formules ci-dessus, on a calculé les taux de réponse des ménages et des femmes dans 66 enquêtes effectuées dans 44 pays entre 1990 et 2000. Les résultats sont présentés dans l'annexe pour les régions suivantes : Afrique subsaharienne, Amérique latine, Asie, Eurasie et Proche-Orient.

29. Les données montrent que les taux de réponse des ménages à ces enquêtes se situent entre 87,9 et 99,5 %, avec une moyenne de 97,5 %, ce qui veut dire que la grande majorité des ménages identifiés dans les échantillons des DHS ont été interrogés avec succès. Pour les mêmes enquêtes, le taux de réponse des femmes s'est situé entre 86,5 et 99,3 %, avec une moyenne de 95 %. On a donc obtenu une entrevue complète de la plupart des femmes admissibles.

30. Hormis en Amérique latine où le taux global de réponse des ménages a été de 95 %, toutes les autres régions ont affiché un taux de réponse moyen des ménages d'environ 98 %. De même que pour les ménages, le taux de réponse moyen des femmes a été plus faible en Amérique latine que dans les autres régions couvertes par le programme DHS : 92 % contre 97 %. À l'intérieur de chaque région, les taux de réponse des ménages et des femmes n'ont guère varié d'un pays à un autre, le coefficient de variation se situant entre 0,4 et 3,7 %.

31. Le taux de réponse moyen des ménages est resté élevé à 97 % durant les trois phases du programme DHS (DH II, DHS III et MEASURE-DHS +)², tandis que le taux de réponse moyen des femmes a augmenté légèrement, passant de 94 à 96 %.

32. Ces taux de réponse élevés, tant au niveau des ménages qu'au niveau des individus sont le résultat d'une formation rigoureuse du personnel de terrain et d'une étroite supervision du travail de terrain. Par ailleurs, lors de chaque enquête, on veille à faire en sorte que le délai entre l'établissement des listes et les entrevues ne soit pas trop long. De plus, à la différence des enquêtes effectuées dans les pays développés, dans les pays en développement, les enquêtes sur les ménages bénéficient généralement d'un haut niveau de coopération de la part des répondants potentiels. Au fil du temps, les taux de réponse moyens des ménages et des individus ont été remarquablement similaires à l'intérieur de chaque région.

H. Effets de conception de l'échantillon des DHS

33. La présente section donne un bref résumé de quelques-uns des effets de conception et des valeurs des coefficients de corrélation intra-classe ($\bar{\rho}$) observés dans les Enquêtes démographiques et sanitaires [voir Lê et Verma (1997) pour plus de détails; Kish, Groves et Krotki (1976) et Verma, Scott et O'Muircheartaigh (1980) pour des analyses similaires des erreurs d'échantillonnage dans les Enquêtes mondiales sur la fécondité (WFS)].

34. L'effet de conception est le ratio de la variance d'échantillonnage de toute estimation obtenue à l'aide d'un échantillon de conception complexe et à la variance

² MEASURE est un vaste projet de l'USAID, dont MEASURE DHS + fait partie. « MEASURE » est le sigle de « Monitoring and Evaluation to Assess and Use Results » (Suivi et évaluation pour mesurer et utiliser les résultats).

de la même estimation appliquée à un échantillon aléatoire simple ou à un échantillon sans restriction de même taille (Kish, 1965), c'est-à-dire :

$$D^2(y) = \frac{Var_{\text{complexe}}(y)}{Var_{\text{sans restriction}}(y)}$$

35. Les effets de conception résultent de la stratification, de probabilités de sélection inégales, d'ajustements des pondérations de l'échantillon (pour les cas de non-réponse), d'ajustements de la pondération de la population (pour non-couverture et pour plus de précision) et de la mise en grappe de tous les éléments d'un échantillon de conception complexe.

36. L'effet de conception estimatif dû à la pondération peut se calculer à partir de l'échantillon au moyen de la formule :

$$d^2(\hat{y}) = 1 + cv^2(w_j)$$

où cv^2 est le carré de la variation du coefficient des pondérations d'échantillonnage w_j .

37. L'effet de conception dû à l'effet de mise en grappe peut se calculer à partir de l'échantillon au moyen de la formule

$$D^2(\hat{y}) = 1 + (b-1)\rho$$

où b est la taille moyenne des grappes et ρ est la corrélation intra-classe.

38. Une analyse détaillée des effets de conception et des coefficients de corrélation intra-classe — définitions, composantes des effets de conception, utilisation des effets de conception et des coefficients de corrélation intra-classe dans la conception des enquêtes par sondage — est présentée au chapitre VI de cette publication. Pour comprendre l'effet d'un échantillon de conception complexe sur les erreurs types, on utilise couramment la racine carrée de l'effet de conception, $d(\hat{y})$.

39. Comme on l'a indiqué précédemment, les enquêtes démographiques et sanitaires sont fondées sur des échantillons de ménages représentatifs à l'échelon national avec conception d'un échantillon probabiliste stratifié à phases multiples standard comprenant un assez grand nombre d'UPE. Les estimations sont généralement établies à l'échelon national, pour les zones urbaines et rurales et pour de petites régions géographiques coïncidant généralement avec les régions administratives de nombreux pays.

40. Lê et Verma (1997) ont étudié les erreurs d'échantillonnage de 48 Enquêtes démographiques et sanitaires effectuées entre 1985 et 1993. Pour les estimations nationales d'ensemble, l'effet de conception moyen $d(\bar{y})$, où \bar{y} était souvent la moyenne d'une proportion établie sur 37 variables et 48 enquêtes, était d'environ 1,50, avec des moyennes allant de 1,13 pour Trinité-et-Tobago à 2,07 pour le Nigéria. Cela veut dire que la mise en grappe, la pondération et d'autres aspects de la conception ont accentué les erreurs types des estimations, en moyenne, par un coefficient de 1,5 (ou les variances des estimations, par un coefficient de 2,25) par rapport à celles d'un échantillon de même taille sans restriction.

41. Dans la plupart des pays, des grappes de tailles similaires ont été utilisées dans les zones urbaines et rurales (grappes de taille moyenne de 24 dans les zones urbaines et de 30 dans les zones rurales). De ce fait, la différence entre les valeurs moyennes de $d(\bar{y})$ entre les zones urbaines et rurales était faible, de 1,4 pour les zones urbaines et de 1,5 pour les zones rurales. On a constaté la même tendance pour les valeurs de $d(\bar{y})$ des différentes régions géographiques. À l'intérieur de chaque pays, les valeurs de $d(\bar{y})$ étaient très proches les unes des autres selon les régions, et n'étaient que marginalement plus faibles que la valeur $d(\bar{y})$ totale du pays, reflétant une fois encore

l'utilisation de la même conception dans toutes les régions du pays. Par contraste, les valeurs de $d(\bar{y})$ étaient sensiblement plus faibles que les valeurs nationales pour les sous-groupes définis par les caractéristiques démographiques et socio-économiques des répondants. Comme ces sous-groupes se retrouvent dans toutes les UPE, les tailles des grappes (b_d) étaient plus petites que les tailles des grappes de l'échantillon total (b), et, par conséquent, les effets de conception des sous-groupes avaient tendance à être plus faibles. Par exemple, dans la DHS effectuée en Tunisie, les valeurs de $d(\bar{y})$ pour la variable « taille idéale de la famille » étaient respectivement de 1,56 et 1,70 pour les sous-groupes de femmes qui travaillent et les sous-groupes de femmes qui ne travaillent pas, alors que la valeur de $d(\bar{y})$ de l'échantillon total était de 1,79.

42. Dans les Enquêtes démographiques et sanitaires, les taux d'échantillonnage différentiels pour les zones urbaines et rurales ou pour les régions géographiques nécessitaient une pondération des données de l'échantillon. La pondération était aussi nécessaire pour compenser les taux différentiels de non-réponse et autres imperfections dans l'application de l'échantillon. Ces pondérations avaient tendance à gonfler les erreurs d'échantillonnage. L'effet de conception dû à des pondérations variables a été calculé pour les Enquêtes démographiques et sanitaires pour des estimations fondées sur les échantillons totaux. Dans les premières enquêtes des années 1985-1990, la valeur moyenne de $d(\bar{y})$ due à la pondération était de 1,08 (représentant un accroissement de 17% de la variance). Cette valeur est passée à 1,15 (représentant un accroissement de 32 % de la variance) dans les enquêtes effectuées dans les années 1990-1993, qui s'écartaient davantage de la coutume d'utiliser des échantillons *epsem* (de probabilité égale) entre les zones urbaines et rurales afin de permettre des estimations régionales.

43. Comme le montre le tableau XXII.1, la moyenne des valeurs de $d(\bar{y})$ de l'échantillon total entre tous les pays varie sensiblement selon les variables, les valeurs de $d(\bar{y})$ allant d'environ 1,1 ou 1,2 pour les variables correspondant à la mortalité infantile, à un niveau élevé de 2,5 pour l'estimation d'une assistance médicale à la naissance. Cela dénote la corrélation plus élevée à l'intérieur des grappes géographiques de la disponibilité de soins de santé. Lors de l'examen de la variabilité de ces valeurs de $d(\bar{y})$, il convient de noter les différences dans les bases d'échantillonnage entre les différentes parties du tableau. Par exemple, l'ensemble d'estimations du haut est fondé sur toutes les femmes âgées de 15 à 49 ans; le second ensemble de données n'est fondé que sur les femmes de cette tranche d'âge qui sont actuellement mariées, et l'ensemble suivant est fondé sur toutes les naissances des cinq dernières années. La variation des bases d'échantillonnage se traduit par des valeurs b différentes dans les effets de conception pour la mise en grappe, et ce facteur contribue à la variabilité des valeurs de $d(\bar{y})$ sur le tableau XXII.1.

44. La mesure de l'homogénéité ρ est plus utile que l'effet de conception en raison de la mise en grappe en vue de prochaines enquêtes, car l'effet de conception dépend à la fois de ρ et de la taille des grappes b . L'effet de conception d'une enquête passée n'est applicable à la nouvelle enquête que si ces deux paramètres sont les mêmes. Cependant, il conviendrait d'envisager de changer b , car la taille des grappes peut être déterminée comme la responsable de l'échantillonnage, ce qui n'est pas le cas de la corrélation intra-classe. Si l'on dispose d'une estimation de ρ , on peut examiner l'effet produit par le changement de b en calculant les effets de conception de la mise en grappe pour différentes valeurs de b . Autrement dit, ρ est le principal facteur à examiner. Des estimations de $\bar{\rho}$ ont été calculées à partir des Enquêtes démographiques et sanitaires, et les résultats sont affichés au tableau XXII.1. Comme le montre ce tableau, les valeurs de $\bar{\rho}$ varient considérablement, allant d'un minimum de 0,01 à un maximum de 0,22. Comme prévu, les estimations qui dépendent de l'existence d'établissements locaux de santé tendent à afficher des valeurs de $\bar{\rho}$ élevées.

Tableau XXII.1
Valeurs moyennes de $d(\bar{y})$ et de $\hat{\rho}$ pour 48 Enquêtes DHS, 1984-1993

Proportion/moyenne	$d(\bar{y})$	$\hat{\rho}$
Toutes les femmes âgées de 15 à 49 ans ^a		
Actuellement mariées	1,43	0,03
Nombre d'enfants mis au monde	1,35	0,02
Nombre de naissances durant les cinq dernières années	1,44	0,03
Nombre d'enfants vivants de moins de cinq ans	1,41	0,02
Nombre d'enfants mis au monde par les femmes âgées de 15 à 49 ans	1,26	0,02
Femmes actuellement mariées âgées de 15 à 49 ans		
Ne voulant plus d'enfants	1,32	0,02
Voulant attendre deux ans ou plus pour leur prochain enfant	1,24	0,01
Connaissant une méthode contraceptive	2,01	0,14
Connaissant une méthode contraceptive moderne	2,08	0,15
Connaissant une source pour l'acquisition de contraceptifs	1,94	0,12
Utilisant actuellement une méthode contraceptive quelconque	1,50	0,05
Utilisant actuellement une méthode contraceptive moderne	1,43	0,04
Utilisant actuellement un dispositif intra-utérin (DIU)	1,42	0,04
Utilisant actuellement la pilule	1,41	0,04
Utilisant actuellement des préservatifs	1,38	0,03
Utilisant actuellement une source publique d'approvisionnement en contraceptifs	1,36	0,03
Stérilisées	1,36	0,03
Toutes les naissances des cinq dernières années		
Assistance médicale à l'accouchement	2,54	0,22
Mère ayant reçu le toroïdal du tétanos	2,02	0,12
Enfants de moins de cinq ans		
Cas de diarrhée au cours des deux dernières semaines	1,34	0,03
Parmi ces cas, enfants ayant été traités aux SROb ^b	1,25	0,12
Enfants âgés de 6 à 35 mois		
Taille inférieure de deux écarts types à la norme pour l'âge	1,33	0,05
Poids inférieur de deux écarts types à la norme pour l'âge	1,29	0,04
Rapport poids/taille de deux écarts types à la norme	1,19	0,02
Enfants âgés de 12 à 23 mois		
Ayant un livret de santé	1,33	0,15
Parmi les enfants ci-dessus, ayant reçu toutes leurs vaccinations	1,31	0,21
Enfant nés il y a 1 à 4 ans ou 5 à 9 ans		
Taux de mortalité infantile des enfants nés 1 à 4 ans avant l'enquête	1,23	0,02
Taux de mortalité infantile des enfants nés 5 à 9 ans avant l'enquête	1,14	0,01

^a Dans environ un quart des enquêtes, l'échantillon et, par conséquent, toutes les variables de ce groupe, ont été limités aux femmes mariées ou ayant été mariées.

^b Sels de réhydratation orale.

45. L'une des importantes constatations qui ressortent des analyses des erreurs d'échantillonnage du programme DHS est que, pour une estimation donnée, les estimations de $\bar{\rho}$ sont relativement transposables d'un pays à un autre, à condition que les échantillons soient comparables dans leur conception. Par conséquent, lors de la

conception d'une nouvelle enquête dans un pays, on peut utiliser les données empiriques sur les erreurs d'échantillonnage d'une enquête similaire dans un pays voisin s'il y a lieu et si l'on prend soin de s'assurer de la comparabilité.

I. Exécution de l'enquête³

46. Alors que l'on accorde une grande attention à l'échantillonnage scientifique et au calcul des erreurs d'échantillonnage, il ne faut pas perdre de vue qu'il existe de nombreuses sources d'erreurs dans les enquêtes. Les erreurs liées à la variabilité de l'échantillonnage peuvent généralement être chiffrées, mais tel n'est généralement pas le cas d'autres types d'erreurs, surtout si l'on n'accorde pas suffisamment d'attention à la formation et au recrutement du personnel de terrain et du personnel de traitement des données. C'est pourquoi le contrôle des erreurs autres que les erreurs d'échantillonnage est l'un des objectifs importants de toute Enquête démographique et sanitaire.

47. Nombre d'Enquêtes démographiques et sanitaires sont effectuées dans des pays où il est difficile de recruter un personnel de terrain hautement qualifié et où le travail de terrain pose de sérieux problèmes de transport, de logement, d'hygiène, d'alimentation, etc. La nécessité pour le personnel de terrain de voyager à travers le pays pose également des problèmes de sécurité et de supervision. Ce sont là certaines des principales raisons pour lesquelles le programme DHS attache beaucoup d'importance à la formation du personnel de terrain et à la supervision sur le terrain et au bureau. Pourtant, malgré cette importance accordée à la supervision, il y a eu des cas où l'exécution a laissé à désirer et où se sont posés des problèmes concernant la qualité des données. On trouvera ci-dessous une description des mesures qui sont prises pour l'exécution d'une enquête démographique et sanitaire et qui soulignent la nécessité d'une préparation détaillée, d'une longue formation et d'une supervision minutieuse.

48. Un autre aspect important des enquêtes a trait à la mesure dans laquelle les données deviennent disponibles en temps voulu et sont accessibles aux décideurs, aux responsables de la gestion des programmes et aux analystes. Trop d'enquêtes, surtout dans les pays en développement, ne sont jamais analysées ou diffusées convenablement. Le programme DHS vise à faire en sorte que toutes les enquêtes fassent l'objet d'une analyse en temps voulu, que leurs résultats soient publiés et diffusés et que les données puissent être utilisées pour des travaux de recherche complémentaires. Le processus nécessaire à cet effet est décrit ci-dessous.

J. Préparation et traduction des documents de l'enquête

49. Les documents utilisés pour l'enquête dans chaque pays participant comprennent généralement un questionnaire pour les ménages, un ou plusieurs questionnaires individuels pour les femmes et/ou les hommes et les manuels correspondants. Les questionnaires comprennent les questions essentielles à toute DHS, certaines questions adaptées au pays et des modules facultatifs. Le personnel de l'enquête travaille avec des homologues locaux à l'adaptation des questionnaires, compte tenu des besoins du pays. Les modèles de questionnaires de DHS sont longs, de sorte que toute addition doit soigneusement tenir compte de la longueur totale du questionnaire. S'il est difficile à traiter et demande trop de temps pour le remplir, la qualité des données risque d'en souffrir. Les versions de base des *Manuels de l'enquêteur et du superviseur* sont adaptées à chaque pays en fonction du contenu du questionnaire propre au pays.

50. La règle du Programme DHS est de veiller à ce que les questionnaires soient traduits et imprimés dans les principales langues locales, de manière à ce que les entre-

³ Une bonne partie des informations figurant dans les sections consacrées à l'organisation de l'enquête et aux caractéristiques des Enquêtes démographiques et sanitaires sont tirées du projet de *Manuel d'organisation des Enquêtes démographiques et sanitaires*, rédigé par l'un des auteurs du présent chapitre.

vues se déroulent dans la langue des répondants. Tout groupe linguistique qui représente 10 % ou plus de l'échantillon doit avoir sa propre version traduite du questionnaire. Souvent, on ne peut éviter complètement que l'enquêteur ou quelqu'un d'autre ait à faire une traduction sur place, car il peut arriver qu'il n'y ait aucune traduction adéquate du questionnaire pour certains répondants. Néanmoins, il faut s'efforcer de limiter le plus possible d'avoir à faire une traduction sur place.

51. La traduction n'est pas chose facile et exige à la fois de solides connaissances linguistiques et la compréhension des termes et expressions typiques des enquêtes démographiques et sanitaires. On trouve rarement toutes ces compétences chez une même personne, surtout lorsque plusieurs langues sont utilisées dans le même pays.

52. Selon l'approche DHS de la traduction, une personne doit traduire le questionnaire DHS dans la langue locale requise en utilisant la version anglaise, espagnole ou française du questionnaire de base. Si un questionnaire a déjà été traduit pour une DHS précédente ou pour une autre enquête de type similaire, cette traduction doit certainement être prise en compte. En règle générale, si les mêmes questions doivent être posées, la traduction devrait être la même, à moins que la traduction précédente n'ait été jugée mauvaise.

53. Le questionnaire traduit est ensuite retraduit dans sa langue d'origine par un traducteur indépendant. Il est essentiel que cette retraduction se fasse sans consulter le questionnaire original, afin d'assurer sa pleine indépendance vis-à-vis de l'original. Ensuite, on réunit les deux traducteurs et l'équipe de direction de l'enquête pour comparer les deux versions et corriger les différences. C'est là un processus important, surtout dans le cas de langues qui ne sont généralement pas écrites, car la traduction n'est alors pas chose aisée.

54. Ce processus devrait déboucher sur des questionnaires compréhensibles des répondants qui sont interrogés dans leur langue. Toutefois, il est également nécessaire de tester les traductions sur le terrain avant d'en généraliser l'application à l'enquête. On n'a pas besoin de procéder à un grand nombre d'entrevues sur le terrain, mais un minimum de trois à cinq devrait être effectué dans chaque langue avant que la traduction soit définitivement adoptée. Il ne faut pas oublier que la traduction a pour but de faire en sorte que la même question soit posée à chaque répondant. Cela ne veut pas dire que la traduction doit être littérale. La bonne traduction est celle qui transmet le même message, même si ce n'est pas mot à mot. Les Enquêtes démographiques et sanitaires sont souvent répétées dans le même pays, mais les questionnaires peuvent différer quelque peu dans leur contenu. On peut donc utiliser d'anciennes traductions de la plupart des questions ainsi que l'expérience acquise lors de précédents cycles de tests et de travaux de terrain.

55. La documentation de l'enquête, en l'occurrence les manuels de l'enquêteur et du superviseur, doit être traduite dans une langue comprise par tout le personnel de terrain si les versions anglaise, espagnole ou française ne peuvent être utilisées.

K. Le test préalable

56. Le test préalable est un moyen crucial de tester les traductions, les questions à sauter dans le questionnaire, les manuels de l'enquêteur et du superviseur et les autres procédures de l'enquête. Il permet également à l'équipe de direction de l'enquête d'acquérir une expérience de la formation de personnel de terrain avant le principal stage de formation. En général, le responsable de l'enquête dans le pays participe aux entrevues effectuées au titre du test préalable.

57. Pour les besoins du test préalable, une formation, qui dure généralement environ deux semaines, est dispensée à un petit nombre d'agents de terrain. Cette formation est dispensée par un personnel local, avec l'assistance du responsable de l'enquête dans le pays. La DHS a pour pratique de former les superviseurs de demain comme enquêteurs pour le test préalable. Ils reçoivent ensuite une formation de superviseurs. Cela permet de leur dispenser une formation très large et de leur assurer déjà un rôle au stade de la formation d'enquêteurs, et cela garantit la présence d'un personnel suffisant pour guider les séances de formation et les tests qui ont lieu pendant la formation des enquêteurs.

58. En général, le test préalable porte sur 100 à 200 ménages, et les entrevues prennent environ une semaine à effectuer. Les entrevues menées dans le cadre du test préalable ont lieu dans des zones urbaines et rurales qui n'ont pas été sélectionnées pour l'enquête principale, de manière à éviter la contamination au niveau des résultats de l'enquête. L'expérience acquise par ce type d'enquête est aujourd'hui très vaste, de sorte que le test préalable peut être limité et n'a plus besoin de couvrir de nombreuses régions du pays.

59. Le travail de terrain nécessaire au test préalable suit les mêmes procédures que celui de l'enquête. Autrement dit, il est établi une liste des ménages, afin que les équipes s'habituent à suivre les procédures et à utiliser leurs formulaires de contrôle. L'équipe de direction de l'enquête supervise activement toutes les phases du test préalable, afin de se familiariser avec les problèmes qui se posent et de pouvoir recommander des solutions.

60. L'expérience du test préalable sert de base à la révision des questionnaires et des manuels de l'enquête. Les erreurs sont corrigées et des améliorations sont apportées sur la base du travail observé pendant cette phase. Il est essentiel à ce stade de tenir un registre de tous les problèmes constatés pendant la formation, les essais et les entrevues proprement dites. Les problèmes observés durant les entrevues font l'objet de rapports établis par l'équipe qui observe les entrevues effectuées dans le cadre du test préalable et de comptes rendus journaliers présentés par les enquêteurs qui participent au test préalable. Il est important que toutes les personnes qui participent à ce test prennent note de ce qu'elles observent.

61. On doit veiller également à ce que toutes révisions effectuées après le test préalable n'introduisent pas de nouvelles erreurs. En fait, s'il faut effectuer une révision approfondie des questionnaires, on procède à quelques entrevues sur le terrain avec le nouveau questionnaire pour s'assurer que les révisions sont bonnes et qu'elles n'ont pas introduit de nouveaux problèmes.

L. Recrutement du personnel de terrain

62. La qualité d'une enquête sur les ménages dépend pour beaucoup de la qualité du personnel de terrain. C'est pourquoi il importe de recruter les meilleurs candidats possibles pour ce travail. Dans les pays en développement, peu d'organisations disposent d'une équipe permanente d'enquêteurs et de superviseurs, et lorsqu'elles disposent d'un tel personnel, il tend à être composé principalement d'hommes. Il faut des enquêteuses pour les enquêtes démographiques et sanitaires, à moins que ces enquêtes ne portent que sur les hommes. C'est ce qui explique que ces enquêtes sont généralement confiées à des personnes spécialement recrutées pour leur exécution. Comme la phase de collecte de données ou de travail de terrain dure généralement de trois à six mois, les personnes recrutées sont le plus souvent des personnes qui n'occupent pas actuellement d'emploi et qui sont prêtes et aptes à passer plusieurs mois loin de chez elles. Dans certains pays où ces enquêtes ont une composante santé plus

importante, c'est un personnel médical travaillant pour le Ministère de la santé qui est détaché en qualité d'enquêteurs et de superviseurs.

63. Le recrutement doit tenir compte du nombre de personnes nécessaires pour parler chacune des langues dans lesquelles doit se dérouler l'enquête. Le nombre de personnes recrutées comme stagiaires est d'au moins 10 à 15 % plus élevé que le nombre de personnes nécessaires pour le travail de terrain, afin de tenir compte des abandons et des rejets de candidats jugés non qualifiés. Le recrutement se fait sur la base d'une évaluation objective des aptitudes des candidats plutôt que de tout autre critère. Les candidats doivent être présentables, capables de parcourir à pied de longues distances et d'établir un bon contact avec les personnes à interroger. Il importe également qu'ils fassent preuve d'un bon esprit d'équipe. En aucun cas le recrutement ne doit être fondé sur les rapports du candidat avec le personnel de l'enquête, sur le favoritisme ou sur toute autre pratique de recrutement inacceptable.

64. Les postes de superviseur et d'éditeur doivent être confiés à des personnes capables d'être des chefs d'équipe. Ces personnes doivent avoir une certaine assurance, une forte motivation et un excellent esprit d'équipe. Toutefois, les principales caractéristiques d'un bon enquêteur sont une aptitude à poser des questions avec facilité et naturel, à mettre le répondant à l'aise et à enregistrer correctement les réponses reçues.

M. Formation des enquêteurs

65. La formation des enquêteurs est très semblable à celle qui est dispensée pour les besoins du test préalable, mais elle dure en général de trois à quatre semaines, en partie du fait du plus grand nombre de participants. Pendant cette formation, les candidats enquêteurs effectuent au moins 10 à 15 entrevues à titre d'essai sur le terrain. La formation est dispensée par un personnel local aidé par le personnel formé pour le test préalable et le responsable de la DHS dans le pays.

66. La sélection finale des enquêteurs se fait sur la base de leurs résultats à une série de tests écrits ainsi que de l'observation de leur comportement durant les entrevues à titre d'essai au bureau et de la qualité de leurs entrevues pour le test préalable. Il est extrêmement important que les critères de sélection soient objectifs. Dans beaucoup d'endroits, l'équipe de l'enquête est soumise à de fortes pressions de la part d'autres personnes pour confier les emplois à des candidats du choix de ces personnes. Or, la seule façon de sélectionner le personnel est de procéder à un examen des qualifications des candidats et à une évaluation objective de leur comportement pendant la formation. En fait, le recours à des tests écrits objectifs pendant la formation peut aider l'équipe de l'enquête à justifier les raisons pour lesquelles certains candidats n'ont pas été acceptés.

N. Travail de terrain

67. La DHS a pour politique d'aborder le travail de terrain comme un travail d'équipe. Il y a de nombreuses raisons à cela, mais la principale est la possibilité d'assurer un plus haut niveau de supervision du travail. Une autre raison est la nécessité de prévoir des moyens de transport spéciaux pour la plupart des enquêteurs. Dans beaucoup de pays, la nécessité de protéger le bien-être du personnel de terrain est une autre raison importante.

68. En général, les équipes sont composées d'un superviseur (le chef d'équipe), d'une éditrice, d'un technicien de santé et des trois à quatre enquêteuses. S'il est également prévu d'interroger des hommes, l'équipe comprend généralement un enquêteur.

Dans la plupart des pays, un véhicule avec chauffeur est affecté à chaque équipe. La taille de l'équipe est parfois limitée par la capacité des véhicules utilisés.

69. Le superviseur est globalement responsable de l'équipe, ainsi que de l'organisation et de la supervision journalières du travail. L'éditrice a pour principale tâche de veiller à la qualité des entrevues. Dans la pratique, le superviseur et l'éditrice doivent partager les mêmes responsabilités afin de former et de maintenir une bonne équipe d'enquêteurs.

70. Les principaux facteurs à prendre en compte lors de la détermination du nombre d'équipes sont le nombre d'UPE, la taille des grappes et la durée prévue du travail de terrain. Toutefois, il importe également de tenir compte du nombre de véhicules disponibles, du nombre d'enquêteurs et de superviseurs pouvant être recrutés et du nombre de langues parlées dans le pays. Le travail de terrain devrait durer de trois à six mois. Il est parfois possible qu'il dure moins longtemps. Toutefois, pour assurer la bonne qualité des données, on limite le nombre d'enquêteurs en raison des difficultés de la formation, de la disponibilité de bons candidats, etc. Cela limite en même temps le nombre d'équipes pouvant être utilisées et détermine la durée du travail de terrain.

71. Si possible, toutes les équipes commencent généralement le travail de terrain dans la même région géographique (par exemple, dans la même province), afin de permettre la supervision de toutes les équipes pour le personnel de direction de l'enquête au moment où cette supervision est la plus nécessaire. Si les équipes se dispersent dans tout le pays dès le début, il est très difficile de leur rendre visite à toutes immédiatement.

72. Les équipes sont affectées à des zones de l'échantillon en fonction des langues parlées et d'autres conditions, ainsi que de la nécessité de faire en sorte que les temps de déplacement de chaque équipe soient maintenus aussi faibles que possible. En général, les équipes travaillent six jours par semaine et doivent rester loin de leurs foyers pendant plusieurs semaines ou plusieurs mois à la fois.

73. Si une entrevue n'est pas terminée à la première visite, on fait de nouveaux essais auprès du ménage ou du répondant sélectionné, jusqu'à trois fois sur trois jours différents, avant de le classer comme cas de non-réponse. Les contacts ultérieurs sont prévus à des heures où le répondant a le plus de chances d'être chez lui. Lorsque la plupart des membres de l'équipe ont terminé leur travail et que de nouvelles visites sont prévues pour un autre jour, il n'est pas rare que l'équipe passe à une nouvelle grappe et laisse un enquêteur en arrière pour finir le travail. Cela est possible lorsque la nouvelle grappe n'est pas trop éloignée et que le véhicule peut prendre l'enquêteur laissé en arrière une fois son travail fini. Comme on l'a indiqué précédemment, il n'est pas prévu de remplacement pour les ménages ou les individus qui refusent de se laisser interroger ou sont de quelque autre manière classés comme cas de non-réponse.

74. Les équipes doivent disposer d'un nombre suffisant de questionnaires et documents afin de ne pas ralentir le rythme du travail. Les questionnaires remplis doivent être emballés, protégés des éléments et sauvegardés jusqu'à leur transmission au siège, généralement par des superviseurs itinérants qui visitent chaque équipe périodiquement.

75. Les Enquêtes démographiques et sanitaires se distinguent par l'importance qu'elles accordent à la supervision. L'expérience montre que, sans une supervision continue, la qualité des données se détériore considérablement. C'est pourquoi il est prévu plusieurs niveaux de supervision. Le superviseur de l'équipe et l'éditeur sont tenus d'observer de temps en temps les enquêteurs et de s'assurer que chaque questionnaire est rempli dans son intégralité et avec exactitude. Lorsqu'ils constatent de gros problèmes, ils demandent aux enquêteurs de retourner voir la personne interrogée

afin de recueillir des informations exactes. De plus, le superviseur est généralement tenu de réinterroger un sous-échantillon d'environ 10 % des ménages sélectionnés afin de s'assurer que l'entrevue initiale a effectivement eu lieu et que toutes les femmes à interroger ont été convenablement identifiées.

76. Le directeur et le personnel de l'enquête DHS assurent une supervision complémentaire pendant le travail de terrain. Ils rendent régulièrement visite aux équipes sur le terrain pour vérifier le travail des enquêteurs, des éditeurs et des superviseurs. Pendant ces vérifications, ils examinent au moins un ou deux questionnaires remplis par chaque enquêteur après leur examen par l'éditeur de l'équipe. Cela permet de repérer en même temps les erreurs de l'enquêteur et de l'éditeur. Les visites de supervision sur le terrain sont extrêmement importantes. Il n'est pas rare que certains superviseurs et éditeurs ne fassent pas un travail vraiment satisfaisant. Cela nuit à la qualité du travail des enquêteurs et doit être redressé le plus tôt possible. Les visites sur le terrain sont le principal moyen de remédier à ce problème. Pendant ces visites, il est utile de recourir aux « tableaux d'évaluation de la qualité des données » qui sont établis à intervalles réguliers pendant le travail de terrain afin de signaler certains problèmes particuliers et les problèmes qui se posent à propos de certaines équipes ou de certains enquêteurs particuliers. Ces tableaux fournissent des informations sur l'âge des répondants et l'âge des jeunes enfants, qui peuvent être utilisées pour vérifier si les répondants ont été convenablement sélectionnés par les enquêteurs. Ils contiennent également des données sur les cas de mortalité infantile et juvénile qui permettent de mesurer le niveau d'omission des décès d'enfants. Ils indiquent également les taux de réponse des ménages et des individus, ce qui permet de mesurer la productivité de chaque équipe et des enquêteurs et de voir si des ménages et/ou des répondants sont délibérément omis de l'enquête. Les problèmes révélés par les tableaux d'évaluation de la qualité des données sont communiqués au personnel de terrain afin qu'ils puissent être évités à l'avenir (voir également section O ci-dessous).

77. L'établissement de la liste des ménages retenus par l'échantillonnage n'est pas décrit dans la présente section sur le travail de terrain. C'est une opération distincte qui est effectuée de deux à trois mois avant le travail de terrain par un personnel spécialisé (voir section E). Le fait de maintenir cette opération séparée du gros du travail de terrain permet d'assurer que l'établissement de cette liste est bien supervisé et que les ménages peuvent être sélectionnés par un personnel qualifié au bureau avant le travail de terrain. En confiant la sélection de l'échantillon au personnel du bureau, on évite les distorsions qui se produisent souvent lorsque les ménages sont sélectionnés par le personnel de terrain, surtout si la personne qui établit la liste est celle qui choisit l'échantillon.

O. Traitement des données

78. Dans les Enquêtes démographiques et sanitaires, le traitement des données commence généralement entre une et deux semaines après le démarrage du travail de terrain et s'achève ordinairement dans le mois qui suit l'achèvement de ce travail. Le personnel chargé d'entrer les données reçoit une formation sur les questionnaires en assistant soit à une partie du stage de formation des enquêteurs, soit à un stage spécial de deux ou trois jours.

79. L'entrée des données se déroule dans une pièce séparée, où le personnel n'est pas dérangé et où les questionnaires sont en sécurité. Cette pièce doit être proche de l'endroit où les questionnaires remplis sont gardés. Tous les questionnaires sont manipulés plusieurs fois pendant l'entrée et l'édition des données, et la proximité entre

le lieu de stockage et la pièce où sont entrées les données peut réduire considérablement la charge de travail et la pression. Le personnel chargé de l'entrée des données ne travaille pas plus de six heures par jour, en raison de la grande attention que requiert leur travail. Selon le nombre d'ordinateurs utilisables pour l'entrée des données, il peut être nécessaire d'utiliser plus d'une équipe pour que le travail d'entrée et d'édition des données puisse se terminer peu après la fin du travail de terrain. On évite si possible de faire appel à deux équipes, car cela peut entraîner des discordances du fait de la multiplicité des superviseurs et des éditeurs de bureau.

80. La règle pour les DHS est d'entrer deux fois les données de tous les questionnaires (« double entrée »), de comparer les résultats et de résoudre les discordances. Une telle vérification à 100 % réduit grandement le volume de réédition nécessaire pour aplanir les incohérences et produire des résultats plus nets et plus précis. La double entrée des données est confiée à deux équipes distinctes, afin d'assurer de meilleurs résultats. Pendant l'entrée des données, chaque questionnaire est soumis à des vérifications des ordres de grandeur, des questions à sauter et des incohérences.

81. L'un des aspects de l'entrée et de l'édition des données est directement lié au contrôle de la qualité des données. Les DHS ont pour règle de produire périodiquement un ensemble choisi de tableaux pendant l'entrée et l'édition des données, afin de dépister les problèmes difficilement identifiables lors de l'entrée et de l'édition manuelles des données figurant dans les questionnaires. Ces « tableaux de vérification sur le terrain » sont destinés à repérer si, par exemple, les enquêteurs manipulent les âges des répondants ou de leurs enfants afin de réduire leur charge de travail en ne signalant pas tous les cas de mortalité infantile et juvénile, ou en entrant un âge inexact au décès. Ces tableaux sont établis une fois qu'un nombre suffisant de questionnaires ont été entrés, disons 300, puis toutes les deux semaines après cela, de sorte que les tendances des répondants à fausser leurs réponses ou leurs caractéristiques puissent être identifiées par l'enquêteur ou l'équipe d'enquêteurs. Des agents de l'organisme chargé de l'exécution de l'enquête ou du personnel de la DHS examinent ces tableaux. Les problèmes sont communiqués aux équipes appropriées, qui peuvent ainsi prendre des mesures pour y remédier.

82. Les tabulations de base qui sont produites pour chaque pays sont celles qui ont été conçues à partir des données recueillies pour le questionnaire de base. Les tabulations qui découlent des questions ajoutées au questionnaire de base sont établies en collaboration avec les personnes/institutions qui ont demandé ces tableaux supplémentaires. Ce travail doit être effectué assez tôt pour que le processus de tabulation se déroule sans heurts. Tous les tableaux sont soigneusement vérifiés, tant par le personnel de la DHS que par ses homologues dans le pays.

83. En raison de la complexité de l'entrée, de l'édition, de l'imputation et de la tabulation des données, ces opérations sont mises au point par le personnel de traitement des données de la DHS, qui visite le pays afin d'y installer les programmes et de mettre en route le processus. En général, le spécialiste du traitement des données retourne dans le pays à la fin du processus de traitement des données afin d'aider à examiner l'ensemble final des données, de recoder certaines variables, d'imputer les données manquantes, d'attacher les coefficients de pondération de l'échantillon et d'appliquer les ensembles de tableaux préalablement désignés pour le rapport préliminaire et le rapport final. Lors de la tabulation des données, les nombres de cas pondérés et non pondérés sont présentés dans les rapports, quoique les calculs utilisent toujours les pondérations finales de l'échantillon.

P. Analyse et rédaction de rapports

84. La base de l'analyse est l'ensemble de tabulations du modèle de DHS modifiées par le responsable de la DHS dans le pays et le personnel du pays hôte pour l'adapter aux questionnaires utilisés. Ces tabulations sont complétées par des tableaux propres au pays qui présentent les données supplémentaires recueillies dans le pays. L'analyse débouche sur un rapport détaillé sur les données de l'enquête.

85. Un bref rapport sur les principales conclusions est également établi de manière à assurer la plus large diffusion possible des données. Ce rapport est établi immédiatement après ou en même temps que le rapport principal et est distribué lors du séminaire national (voir section Q ci-dessous sur la diffusion).

86. Outre la production de ces rapports d'enquête, la DHS aide les pays à effectuer une analyse plus approfondie des données de l'enquête. Cette analyse débouche généralement sur un document de recherche de 30 à 60 pages, traitant de sujets intéressant particulièrement le pays ou les organismes de financement, mais elle peut également être constituée de tabulations spéciales et de brefs exposés analytiques de nature à permettre au pays de répondre aux questions de politiques et/ou autres questions pertinentes.

Q. Diffusion

87. La diffusion des résultats de l'enquête à tous les publics intéressés est l'un des objectifs essentiels du programme. Les rapports d'enquête font l'objet d'une large diffusion à l'échelon local et sont également distribués aux organismes de coopération et autres institutions présentes dans le pays. Ils peuvent également être consultés et téléchargés à partir du site web de la DHS. Des tableaux muraux, des livres de tableaux, des calendriers, des affiches et autres matériels sont également produits à l'occasion du séminaire national afin d'assurer une plus ample diffusion des résultats de l'enquête.

88. Par ailleurs, un séminaire national est organisé pour présenter les principales conclusions de l'enquête aux décideurs, aux responsables de programmes, aux chercheurs et aux représentants des organisations d'aide. Ce séminaire est généralement couvert par les médias, ce qui favorise l'utilisation des données pour l'élaboration des politiques et des programmes. Certains pays organisent des séminaires régionaux pour faire en sorte que les résultats soient connus et utilisés par-delà les politiques et les programmes nationaux.

89. Toutes les données de ces enquêtes sont entrées dans les archives des DHS. Presque tous les pays qui participent à ce programme ont autorisé l'utilisation de leurs données par les chercheurs responsables du monde entier. L'équipe chargée des archives des DHS suit les demandes de données et fournit les données et la documentation demandées aux personnes habilitées à les utiliser. Ces données sont à présent disponibles gratuitement sur l'Internet, après inscription électronique et autorisation de chaque utilisateur. À la fin de 2002, ORC Macro avait donné accès aux fichiers et sous-fichiers de données DHS plus de 80 000 fois. L'adresse du site Web est : *www.measuredhs.com*. Ce site Web permet également d'obtenir des renseignements complémentaires sur le Programme DHS.

R. Utilisation des données DHS

90. Les données DHS sont généralement utilisées pour suivre et évaluer les progrès des programmes de santé maternelle et infantile et des programmes de population des pays participants. La possibilité d'obtenir des informations sur plusieurs

enquêtes successives permet aux pays de suivre les tendances qui se dégagent des données et de mesurer leurs progrès. Les données sont parfois utilisées pour des programmes d'action immédiate comportant, par exemple, l'administration de suppléments alimentaires à base de fer dans les régions où l'anémie est générale. Le plus souvent, elles servent à façonner la politique et à modifier les objectifs des programmes d'intervention, ainsi que pour la planification à long terme des programmes de santé et de population. Les données des enquêtes DHS ont contribué de façon déterminante à mobiliser l'appui aux programmes de planification familiale en Afrique subsaharienne et ailleurs en montrant que le changement est possible et se produit même dans certains des pays les plus pauvres.

S. Capacités statistiques

91. L'un des buts du Programme DHS est d'accroître l'aptitude des pays participants à recueillir et analyser des données par de vastes enquêtes nationales sur les ménages. Les principaux mécanismes utilisés sont l'établissement d'une documentation de base, telle que questionnaires et manuels, la conception de logiciels qui facilitent le déroulement de l'enquête dans les pays en développement, et la formation sur le tas d'homologues locaux durant toutes les phases de l'enquête.

92. La conception d'un nouveau logiciel apporte une contribution essentielle au renforcement des capacités. Au départ, la DHS a élaboré un programme de traitement des enquêtes appelé Système intégré pour l'analyse des enquêtes [*Integrated System for Survey Analysis* (ISSA)]. L'existence de ce logiciel a été déterminante pour l'obtention rapide de fichiers de données et de rapports. Pour s'adapter aux nouveaux progrès des matériels et des logiciels, la DHS a lancé un nouveau logiciel de traitement des données appelé *Census and Survey Processing* (CSPRO) [Traitement des données des recensements et enquêtes], en collaboration avec le United States Bureau of the Census et une entreprise de conception de logiciels. On prévoit que ce logiciel sera très largement utilisé par différentes institutions pour le traitement des grandes enquêtes. Le United States Bureau of the Census appuie d'ores et déjà de vastes programmes de formation pour l'utilisation de ce logiciel, et l'on prévoit que celui-ci deviendra la norme dans la plupart des pays, ce qui facilitera grandement les efforts de renforcement des capacités.

93. Le programme DHS dispense régulièrement une formation et une information aux homologues locaux à travers une documentation de base détaillée pour l'exécution des enquêtes, des visites régulières au titre de l'assistance technique (10 à 14 par pays) et une coopération pour la préparation des rapports d'enquête. La documentation de base comprend des manuels sur toutes les phases importantes du déroulement de l'enquête. Ces trois mécanismes restent les principaux moyens de renforcer les capacités dans les pays participants.

T. Enseignements

94. La DHS et les enquêtes qui l'ont précédée ont permis de tirer de nombreuses leçons fort utiles pour les enquêtes sur les ménages dans les pays en développement, par exemple :

- Dans beaucoup de pays, les cadres d'échantillonnage ont besoin de coûteuses mises à jour sur le terrain pour pouvoir être utilisés pour les enquêtes nécessitant la collecte de données de qualité. Les listes de ménages sont souvent vieilles et parfois totalement inexistantes. Le contrôle de qualité oblige à sélectionner les ménages au bureau plutôt que de laisser cette sélection au per-

sonnel de terrain, de manière que tous les ménages aient une probabilité de sélection connue. La sélection des ménages au bureau élimine les problèmes causés par la tendance des enquêteurs à visiter les ménages qui sont plus accessibles et à laisser de côté ceux qui sont plus éloignés. La sélection effectuée à partir d'une liste de ménages au bureau permet une sélection objective des ménages sur la liste et facilite la supervision de la sélection de l'échantillon sur le terrain.

- L'actualisation de l'échantillon, une fois effectuée à l'avant-dernière phase de l'échantillonnage, doit être étroitement supervisée de manière à permettre l'établissement d'une liste complète de tous les ménages. On a également remarqué dans un certain nombre d'enquêtes que les responsables de l'établissement de la liste des ménages peuvent être tentés de laisser de côté les habitations qui sont plus reculées ou qui sont situées dans des endroits difficiles d'accès ou dangereux. Sans une bonne supervision, la liste ainsi établie peut être faussée.
- Les taux de réponse sont généralement très bons, tant au niveau des ménages qu'à celui des individus (voir section F sur les taux de réponse).
- Il faut calculer les erreurs d'échantillonnage et les effets de conception sur un ensemble représentatif d'éléments de l'enquête afin d'évaluer l'efficacité de la conception de l'échantillon et la précision des estimations de l'enquête.
- La taille optimale des grappes est de 15 à 20 femmes pour les Enquêtes démographiques et sanitaires, où il faut concilier divers éléments démographiques et sanitaires, certains plus agglutinés que d'autres, et certains portant sur les enfants en bas âge des femmes sélectionnées, et tenir compte du coût de collecte des données.
- L'effet de conception dû à la mise en grappe est de plus en plus fonction de la taille des grappes b et du coefficient de corrélation intra-classe ρ . Comme ρ est relativement transposable entre pays dont les échantillons sont de conception comparable, ρ , b et les effets de conception d'une enquête peuvent être utilisés pour concevoir une enquête comparable dans un autre pays, comme il est indiqué au chapitre VI.
- Pour les enquêtes complexes, la formation d'enquêteurs et de superviseurs dure de trois à quatre semaines. En général, pour les DHS, elle dure trois semaines. Cependant, souvent, cette formation a dû être prolongée d'une semaine ou plus pour assurer la bonne préparation du personnel de terrain. La plupart des problèmes qui se posent sont dus au personnel de terrain, et non pas aux répondants. Une bonne formation et une bonne supervision sont les principaux moyens d'éviter ces problèmes.
- Les enquêteurs et les superviseurs peuvent être la cause de sérieux problèmes. Il est donc nécessaire que la supervision et le contrôle de qualité soient continus, de manière à éviter un travail bâclé et/ou une manipulation délibérée de l'échantillon ou de l'entrevue par certains enquêteurs ou superviseurs cherchant à alléger leur charge de travail. Les enquêtes DHS ont amplement montré que les enquêteurs ont tendance à classer certaines femmes et/ou certains enfants en dehors des tranches d'âge admissibles afin de ne pas avoir à les interroger. Bien que ce problème ne soit pas général à tout le personnel de terrain, il n'en existe pas moins et se limite souvent à quelques-unes des équipes d'enquêteurs. Il faut donc faire preuve d'une vigilance constante durant tout le travail de terrain.
- Il faut s'efforcer de limiter la durée maximum des entrevues en moyenne à une heure. Cette recommandation est fondée non pas sur l'expérimentation

effective sur le terrain de différentes durées, mais plutôt sur l'expérience du personnel de terrain. Les Enquêtes démographiques et sanitaires sont de longueur très variable selon les caractéristiques des répondants et la facilité avec laquelle ils peuvent se rappeler certaines dates et certains événements. La durée d'une entrevue peut aller de 10 minutes pour une femme célibataire sans enfants et sans activité sexuelle à plus d'une heure et demie pour une femme ayant un grand nombre d'enfants, qui ne se rappelle pas facilement les événements qui forment la substance de l'enquête.

- L'un des principaux obstacles en ce qui concerne la logistique sur le terrain a trait à la possibilité de disposer de véhicules adéquats pour le transport des équipes d'enquêteurs. Les véhicules servant au travail de terrain sont coûteux à acheter et à utiliser car ils nécessitent les caractéristiques d'un véhicule tout-terrain pour transporter toute l'équipe. L'absence de véhicules appropriés se traduit par une perte de temps et sape le moral de l'équipe. Même avec de bons véhicules, les enquêteurs et les superviseurs doivent encore parcourir de longues distances à pied pour se rendre dans certaines demeures. Par conséquent, il importe de faciliter le plus possible leur transport jusqu'au lieu de l'enquête.
- L'un des aspects les plus difficiles de la logistique de terrain est de réunir à la fois le bon enquêteur et le bon répondant et d'utiliser le bon questionnaire dans les pays où l'entrevue se fait dans plusieurs langues. La composition des équipes en fonction des aptitudes linguistiques, combinée à un plan de déploiement détaillé tenant compte des qualifications linguistiques des équipes est nécessaire pour faire en sorte que la plupart des répondants soient interrogés dans leur langue vernaculaire par un enquêteur qui parle cette langue et utilise un questionnaire dans cette langue.
- Le personnel chargé de l'entrée des données doit suivre le stage de formation des enquêteurs pour pouvoir entrer les données et les éditer. Les questionnaires DHS sont très compliqués. La participation à la formation des enquêteurs donne à ce personnel une bonne compréhension de l'articulation du questionnaire et de la façon dont les différentes parties de celui-ci sont liées entre elles. Cette compréhension leur est nécessaire pour pouvoir apporter les corrections voulues durant le processus interactif d'entrée et d'édition des données.
- La double entrée des données, si elle peut paraître coûteuse, permet de gagner du temps à l'édition. Lors des premières Enquêtes démographiques et sanitaires, les données n'étaient entrées qu'une fois. Les enquêtes plus récentes utilisent la double entrée afin de déceler les erreurs qui n'ont pu être décelées lors des programmes de vérification de cohérence et pour faire en sorte qu'un minimum de questionnaires ait besoin de corrections au stade de l'édition. Le personnel chargé du traitement des données des enquêtes DHS a conclu que l'effet bénéfique de la double entrée des données sur leur édition l'emporte largement sur son coût.
- Il importe d'informer constamment le personnel de terrain des problèmes constatés lors de l'entrée des données au sujet des questionnaires remplis afin d'assurer l'excellente qualité des données. Le personnel de terrain a besoin, surtout au début de l'enquête, d'être informé immédiatement des erreurs qu'il a commises, afin d'éviter qu'il les répète par la suite. L'interaction au stade de l'entrée des données est un excellent moyen de dépistage précoce des problèmes qui se posent sur le terrain.
- Il importe d'examiner certains tableaux afin de déceler des tendances dans les réponses qui n'apparaissent pas nécessairement à l'édition des questionnaires.

naires. Par exemple, les enquêteurs classent-ils délibérément d'éventuels répondants comme plus âgés ou plus jeunes afin d'éviter d'avoir à les interroger ? Ce type de problèmes ne peut être clairement identifié que par l'étude des profils d'âge des répondants à plusieurs centaines d'entrevues.

- Dans beaucoup de pays, l'établissement du rapport d'enquête est l'une des tâches les plus difficiles. Le renforcement des capacités en étude d'enquêtes est l'un des objectifs du programme DHS. La rédaction de rapports est l'un des domaines où un gros effort est consacré à renforcer les capacités par un travail interactif avec des rédacteurs locaux. Depuis quelque temps, les ateliers d'étude de la rédaction de rapports au cours desquels tous les rédacteurs travaillent sur certains chapitres de ces rapports en collaboration avec le personnel DHS sont considérés comme l'un des moyens les plus efficaces de transfert de capacités. Néanmoins, la rédaction de rapports est aussi un art auquel tous, quel que soit leur niveau de connaissance de la démographie, ne présentent pas les mêmes prédispositions.
- L'assistance technique est particulièrement nécessaire aux stades de l'échantillonnage, du traitement des données et de la rédaction du rapport. Dans les autres domaines, cette assistance vise souvent à faire en sorte que les différentes opérations soient exécutées en bon ordre chronologique. Les domaines mentionnés précédemment sont ceux qui posent le plus de difficultés au personnel local de beaucoup d'Enquêtes démographiques et sanitaires, sinon de toutes. Par comparaison, nombre d'organismes locaux assurent très bien la formation et le travail de terrain. Il importe donc de cibler l'assistance technique sur les domaines où les faiblesses ont besoin d'être surmontées.
- Les pays sont prêts à partager les données de leurs enquêtes avec des chercheurs responsables. Ces questions doivent être réglées avant l'exécution de l'enquête. Le Programme d'enquêtes démographiques et sanitaires a très bien réussi à obtenir l'approbation des pays participants en ce qui concerne le partage de leurs données en vue de futurs travaux de recherche. Cela a permis de créer une base de données portant sur plusieurs pays, qui est devenue un outil d'une valeur inestimable tant pour les pays que pour les donateurs. Pour atteindre cet objectif, il a fallu conclure des accords avec les autorités des pays participants au moment de l'enquête. Lorsqu'un tel accord ne peut être conclu à ce stade, il peut souvent être négocié à un stade ultérieur par suite d'un changement de gouvernement ou des personnes responsables en place au moment où l'enquête avait été initialement envisagée.

Annexe

Taux de réponse des ménages et des femmes à 66 enquêtes effectuées dans 44 pays, 1990-2000, de différentes régions

Région	Pays	Année de l'enquête	Phase	Nombre de ménages	Taux de réponse des ménages (%)	Nombre de femmes	Taux de réponse des femmes (%)
Afrique subsaharienne	Bénin	1996	DHS III	4 562	98,6	5 719	96,0
	Burkina Faso	1992	DHS II	5 283	97,3	6 848	92,8
	Burkina Faso	1999	DHS III	4 871	98,8	6 740	95,6
	Cameroun	1991	DHS II	3 647	97,0	4 147	93,3
	Cameroun	1998	DHS III	4 791	98,0	5 760	95,5
	République centrafricaine	1994	DHS III	5 583	99,4	6 005	98,0
	Tchad	1997	DHS III	6 930	98,7	7 705	96,7
	Comores	1996	DHS III	2 277	98,9	3 160	96,5
	Côte d'Ivoire	1994	DHS III	5 977	99,3	8 271	97,9
	Ghana	1993	DHS III	5 919	98,4	4 700	97,1
	Ghana	1999	MEASURE	6 055	99,1	4 970	97,4
	Guinée	1999	MEASURE	5 216	97,6	7 117	94,9
	Kenya	1993	DHS III	8 185	97,1	7 952	94,8
	Kenya	1998	DHS III	8 661	96,8	8 233	95,7
	Madagascar	1992	DHS II	6 027	98,6	6 520	96,0
	Madagascar	1997	DHS III	7 349	97,6	7 424	95,1
	Malawi	1992	DHS II	5 409	98,4	5 020	96,6
	Mali	1996	DHS III	8 833	98,7	10 096	96,1
	Mozambique	1997	DHS III	9 681	95,9	9 590	91,5
	Namibie	1992	DHS II	4 427	92,6	5 847	92,7
	Niger	1992	DHS II	5 310	98,7	6 750	96,3
	Niger	1997	DHS III	6 007	98,7	7 863	96,4
	Nigéria	1990	DHS II	9 173	98,1	9 200	95,4
	Nigéria	1999	MEASURE	7 736	98,8	10 529	93,2
	Rwanda	1992	DHS II	6 292	99,4	6 947	94,3
	Sénégal	1993	DHS II	3 563	99,0	6 639	95,0
	Sénégal	1997	DHS III	4 855	98,3	9 186	93,5
	Togo	1998	DHS III	7 620	98,6	8 964	95,6
	Ouganda	1995	DHS III	7 671	98,4	7 377	95,8
	République-Unie de Tanzanie	1992	DHS II	8 560	97,3	9 647	95,8
	République-Unie de Tanzanie	1996	DHS III	8 141	97,9	8 501	95,5
	Zambie	1992	DHS II	6 245	99,4	7 247	97,4
	Zambie	1996	DHS III	7 365	98,9	8 298	96,7
	Zimbabwe	1994	DHS III	6 075	98,5	6 408	95,6
	Zimbabwe	1999	MEASURE	6 512	97,8	6 208	95,2

Région	Pays	Année de l'enquête	Phase	Nombre de ménages	Taux de réponse des ménages (%)	Nombre de femmes	Taux de réponse des femmes (%)
Amérique latine	Bolivie	1994	DHS III	9 335	97,6	9 316	92,3
	Bolivie	1997	DHS III	12 281	98,6	1 831	94,6
	Brésil	1991	DHS II	6 416	94,5	6 864	90,7
	Brésil	1996	DHS III	14 252	93,2	4 579	86,5
	Colombie	1990	DHS II	8 106	91,4	9 715	89,0
	Colombie	1995	DHS III	11 297	89,5	2 086	92,2
	Colombie	2000	MEASURE	11 747	92,8	2 531	92,5
	République dominicaine	1991	DHS II	8 131	87,9	8 200	89,3
	République dominicaine	1996	DHS III	9 026	97,8	9 034	93,2
	Guatemala	1995	DHS III	11 754	96,1	3 388	92,6
	Haïti	1994	DHS III	4 944	97,5	5 709	93,8
	Nicaragua	1997	DHS III	11 726	98,3	4 807	92,1
	Paraguay	1990	DHS II	5 888	96,5	6 262	93,1
	Pérou	1992	DHS II	13 711	98,3	17 149	92,6
Asie	Bangladesh	1994	DHS III	9 255	99,1	9 900	97,4
	Bangladesh	1997	DHS III	8 762	99,1	9 335	97,8
	Indonésie	1991	DHS II	27 106	99,1	23 470	97,6
	Indonésie	1997	DHS III	34 656	98,8	29 317	98,3
	Pakistan	1991	DHS II	7 404	97,2	6 910	95,7
	Philippines	1993	DHS III	13 065	99,5	15 332	98,0
	Philippines	1998	DHS III	12 567	98,7	14 390	97,2
Eurasie	Kazakhstan	1995	DHS III	4 232	98,7	3 899	96,7
	Kazakhstan	1999	MEASURE	5 960	98,1	4 906	97,8
	Kirghizistan	1997	DHS III	3 695	99,4	3 954	97,3
	Turquie	1993	DHS III	8 900	96,8	6 862	95,0
	Turquie	1998	MEASURE	8 596	93,8	9 468	90,6
	Ouzbékistan	1996	DHS III	3 763	98,4	4 544	97,2
Proche-Orient	Égypte	1992	DHS II	10 950	98,3	9 978	98,9
	Égypte	1995	DHS III	15 689	99,2	14 879	99,3
	Maroc	1992	DHS II	6 635	99,1	9 587	96,5
	Yémen	1991	DHS II	12 934	99,2	6 515	92,2

REMERCIEMENTS

Les auteurs remercient de leurs précieux commentaires les membres d'un groupe d'experts réuni par l'Organisation des Nations Unies pour examiner le texte de cette publication, les réviseurs extérieurs et M. Alfredo Aliaga d'ORC Macro.

RÉFÉRENCES

- Cleland, J., et C. Scott, eds. (1987). *The World Fertility Survey. An Assessment*. New York, Oxford University Press
- Institute for Resource Development/Macro Systems, Inc. (1990). *An Assessment of DHS-I Data Quality*. Methodological Report, No. 1. Columbia, Maryland.
- _____ (1994). *An Assessment of the Quality of Health Data in DHS-I Surveys*. Methodological Report, No. 2. Columbia, Maryland. *Review*, 56(3) : 259-78.
- Kish, L. (1965). *Survey Sampling*. New York, Wiley.
- _____, R. Groves et K. Krotki (1976). *Sampling Errors in Fertility Surveys*. World Fertility Survey Occasional Paper, No. 17, La Haye, International Statistical Institute.
- Lê, T., et V. Verma (1997). *An Analysis of Sample Designs and Sampling Errors of the Demographic and Health Surveys*. DHS Analytical Reports, No. 3. Calverton, Maryland, Macro International, Inc.
- ORC Macro (1996). *Sampling Manual*. DHS-III Basic Documentation, No. 6. Calverton, Maryland.
- _____ (2001). Survey organizational manual. Draft. Calverton, Maryland.
- Scott, Christopher, *et al.* (1988). Verbatim questionnaires versus field translation or schedules: an experimental study. *International Statistical Review*, vol. 56, n° 3, pp. 259-278.
- Vaessen, Martin, *et al.* (1987). Translation of questionnaires into local languages. Dans *The World Fertility Survey: An Assessment*. John Cleland et Chris Scott, eds. New York, Oxford University Press.
- Verma, V., et T. Lê (1996). An analysis of sampling errors for the Demographic and Health Surveys. *International Statistical Review*, vol. 64, pp. 265-294.
- _____, C. Scott et C. O'Muircheartaigh (1980). Sample designs and sampling errors for the World Fertility Survey. *Journal of the Royal Statistical Society A*, vol. 143, pp. 431-473.

Chapitre XXIII

Enquêtes pour l'étude sur la mesure des niveaux de vie

KINNON SCOTT
Banque mondiale
Washington, D.C.
États-Unis d'Amérique

Diane Steele
Banque mondiale
Washington, D.C.
États-Unis d'Amérique

Tilahun Temesgen
Banque mondiale
Washington, D.C.
États-Unis d'Amérique

RÉSUMÉ

Le programme d'études sur la mesure des niveaux de vie [*Living Standards Measurement Study (LSMS) programme*] est né de la nécessité d'améliorer les données statistiques au niveau des ménages de manière à pouvoir concevoir, mettre en œuvre et évaluer les politiques sociales et économiques des pays en développement. Le but visé par le programme LSMS a été de comprendre, de mesurer et de suivre les conditions de vie, l'interaction entre les dépenses et les programmes publics et le comportement des ménages, de permettre des évaluations *ex ante* et *ex post* des politiques et d'analyser les causes des phénomènes observés dans les secteurs sociaux. Les enquêtes LSMS ainsi réalisées utilisent de multiples instruments pour obtenir les données nécessaires à ces fins et s'appuient sur d'importants mécanismes de contrôle pour assurer la qualité des données obtenues. Ces dernières années, en particulier, le Programme LSMS a mis l'accent sur la participation des utilisateurs des données à la conception des enquêtes et sur les questions de durabilité. Le présent chapitre donne un aperçu général de ce que sont les enquêtes LSMS et des principales méthodes de conception et d'exécution de ces enquêtes, ainsi que les efforts en vue de promouvoir la capacité d'analyse. Il offre une évaluation des coûts de l'enquête et de la qualité des données recueillies. Il traite également des calculs des effets moyens de conception de l'échantillon et des coefficients de corrélation intra-classe de certaines variables concernant les ménages et les individus à partir de certaines enquêtes LSMS.

Termes clés : mesure de la pauvreté, niveaux de vie, méthodologie de l'enquête, effet de conception, corrélation intra-classe, contrôle de qualité.

A. Introduction

1. Les dépenses consacrées par le secteur public aux services sociaux et à l'infrastructure représentent un important volume de ressources, tant en valeur absolue que relative. Il n'est pas rare que les dépenses de santé et d'éducation représentent chacune entre 3 et 4 % du produit intérieur brut (PIB). Selon le pays, cela peut se chiffrer à plusieurs millions, voire à plusieurs centaines de millions de dollars. De profondes modifications des politiques fiscales et des politiques des prix influent sensiblement sur les niveaux de bien-être tant absolus que relatifs. Or, souvent, en raison du manque de données, les politiques sont conçues, mises en œuvre et révisées sans grande connaissance de leur efficacité globale sur l'amélioration de la vie des populations du pays. L'absence de données appropriées au niveau des ménages force les décideurs à

s'appuyer sur des données administratives qui, satisfaisantes à certaines fins, limitent souvent très fortement l'aptitude à comprendre le comportement des ménages, la façon dont les politiques gouvernementales influent sur les ménages et les individus et sur les déterminants des résultats observés dans les secteurs sociaux. Les enquêtes sur les ménages ont précisément pour rôle de combler ces écarts de compréhension.

2. Le Programme d'études sur la mesure des niveaux de vie (LSMS) est un instrument que les gouvernements peuvent utiliser, et utilisent, pour mieux comprendre les causes des résultats observés ainsi que l'impact de leurs politiques. L'enquête LSMS ne se contente pas de mesurer les résultats; elle permet d'établir des liens de cause à effet avec ces résultats. Les enquêtes sur les ménages à un seul thème permettent de recueillir des informations approfondies sur un thème particulier, mais ne suffisent pas pour expliquer les raisons de certains résultats et les divers facteurs dont ils sont issus. L'enquête LSMS a pour objet de rechercher les liens entre, d'une part, les divers actifs et les caractéristiques du ménage et, d'autre part, les décisions des pouvoirs publics, et de comprendre ainsi les forces qui agissent sur chaque secteur, sur certains comportements et sur les résultats. En aidant le gouvernement à mieux comprendre les facteurs qui influent sur les conditions de vie, elle lui permet d'améliorer ses politiques et ses programmes, ce qui débouche à son tour sur une utilisation plus efficace et efficiente des maigres ressources publiques et privées et sur une amélioration des niveaux de vie.

3. Les enquêtes LSMS représentent un travail de collaboration entre les gouvernements qui les administrent, les principaux utilisateurs des données dans les pays et la Banque mondiale, ainsi que d'autres institutions d'aide bilatérale et multilatérale¹. Bien que fondée sur des concepts fondamentaux, l'enquête LSMS est largement adaptée pour répondre aux besoins spécifiques de chaque gouvernement à un moment donné. L'institution principale chargée de la mise en œuvre est habituellement l'Office national de statistique [*National statistical office* (NSO)] qui décide de la conception du questionnaire et de l'échantillonnage ainsi que de la méthodologie de terrain utilisant les techniques estimées les plus efficaces pour la LSMS.

4. Le présent chapitre offre un aperçu de l'Étude sur la mesure des niveaux de vie. Il donne d'abord un bref historique du programme, suivi d'informations sur les principales caractéristiques de l'enquête LSMS. Il consacre ensuite une section à expliquer comment les caractéristiques de conception de l'enquête LSMS influent sur la qualité des données recueillies. Dans la dernière section sont présentés des exemples de façons dont les données de ces enquêtes ont été utilisées.

B. Pourquoi une enquête LSMS ?

5. Les enquêtes LSMS ont débuté en 1980 pour répondre au besoin des décideurs de disposer de données de qualité. Au terme d'une période de cinq ans au cours de laquelle les enquêtes pratiquées alors sur les ménages ont été examinées et de longues consultations ont eu lieu avec les chercheurs et les décideurs afin de déterminer les types de données nécessaires, ainsi qu'avec les spécialistes de la méthodologie des enquêtes sur les moyens de concevoir les meilleures procédures de travail de terrain, en 1985, les premières enquêtes LSMS ont été lancées à titre pilote en Côte d'Ivoire et au Pérou. Ces deux premières enquêtes étaient en fait des projets de recherche destinés à tester la méthodologie dans son ensemble afin de déterminer l'utilité et la qualité des données qui ont pu être recueillies². Le succès de ces deux premières enquêtes est à l'origine de plus de 60 enquêtes LSMS effectuées dans plus de 40 pays depuis 1985 (voir la liste complète de ces enquêtes à l'annexe I).

¹ Parmi les autres institutions ayant participé aux enquêtes LSMS, on peut citer notamment la Banque interaméricaine de développement, certaines institutions spécialisées de l'ONU telles que le Programme des Nations Unies pour le développement, le Fonds des Nations Unies pour l'enfance et certains donateurs bilatéraux : Canada, Danemark, États-Unis d'Amérique, Japon, Norvège, Royaume-Uni de Grande-Bretagne et d'Irlande du Nord et Suède.

² Pour un rappel plus détaillé de l'historique du programme LSMS, voir Grosh et Glewwe (1995).

C. Principales caractéristiques des enquêtes LSMS

6. On trouvera ci-dessous un résumé des principales caractéristiques des enquêtes LSMS. Le lecteur est invité à se reporter au Manuel LSMS pour des informations plus détaillées sur ces enquêtes et la façon de les mener à bien³.

³ Dans Grosh et Muñoz (1996).

1. Contenu et instruments utilisés

7. Les enquêtes LSMS utilisent jusqu'à quatre instruments distincts : *a*) un questionnaire à remplir par les ménages pour recueillir des informations aux niveaux des ménages et des individus, ainsi que sur les activités économiques des ménages (agriculture et petit commerce); *b*) un questionnaire sur la communauté⁴, afin de recueillir des données sur le milieu dans lequel fonctionnent les ménages, et plus particulièrement sur les services qui leur sont offerts, les activités économiques, l'accès aux marchés et, depuis peu, le capital social; *c*) un questionnaire sur les prix administré dans chaque région où se trouvent les ménages, afin de permettre des ajustements en fonction du coût de la vie⁵; et *d*) des questionnaires sur les services à remplir par les prestataires locaux de services, afin de recueillir des informations sur les types et la qualité des services offerts aux ménages. La figure XXIII.1 établit un rapport entre les instruments utilisés et les finalités politiques des enquêtes LSMS et les variables nécessaires.

⁴ Il convient de noter qu'il ne s'agit pas de la « communauté » au sens sociologique du terme mais plutôt d'un mécanisme de collecte de données sur les régions où se trouvent les ménages sélectionnés pour l'enquête.

⁵ Les indices nationaux des prix à la consommation sont souvent insuffisants à cet égard, car ils tendent à reposer sur le milieu urbain, et même quand ils englobent les zones rurales, les prix ne sont pas saisis au niveau de fragmentation approprié.

Figure XXIII.1

Rapport entre l'objet des enquêtes LSMS et les instruments de l'enquête

Objet	Indicateurs	Instruments
Mesure du bien-être au niveau des individus et des ménages		
Niveaux, distribution et corrélats	Consommation Revenu Richesse, épargne Capital humain Anthropométriques	Questionnaire pour les ménages Questionnaire sur les prix
Analyse des politiques		
À qui bénéficient les programmes/dépenses publiques	Utilisation des services Qui bénéficie des services, des transferts	
Impact des programmes/dépenses publiques	Coûts des services Impact des politiques	Questionnaire pour les ménages Questionnaire sur la communauté
Disponibilité des services	Distance par rapport au service le plus proche	Questionnaire sur les services Questionnaire sur les prix
Qualité des services	Types de services offerts	
Prix des services	Personnel, budget, autres éléments	
Effet des politiques économiques	Transferts nets entre secteurs	
Identification des déterminants		
Pourquoi les résultats observés se produisent	Composition des ménages	Questionnaire pour les ménages Questionnaire sur la communauté
Ce qui influe sur le comportement des ménages	Capital humain, bien-être, services offerts, etc.	Questionnaire sur les services Questionnaire sur les prix

8. Le contenu des instruments de l'enquête est fonction des besoins prioritaires du pays en matière de données à un moment donné. Comme le souci majeur est de mesurer les niveaux de vie sous tous leurs aspects, l'instrument de l'enquête sur les ménages, en particulier, tend à recueillir des informations sur le vaste ensemble de sujets qui influent sur ces niveaux. Le tableau XXIII.1 illustre le contenu d'une enquête LSMS type, en l'occurrence celle de 1997-1998 au Viet Nam.

Tableau XXIII.1

Contenu du questionnaire de l'enquête sur les ménages au Viet Nam, 1997-1998

Première visite	Seconde visite
Liste des ménages	Fécondité
Éducation	Agriculture, forêts et pêche
Santé	Emploi indépendant non agricole
Main-d'œuvre	Dépenses alimentaires et production
Migration	Produits non alimentaires et biens durables
Logement et services d'utilité publique	Revenu provenant des envois de fonds de travailleurs émigrés Emprunts, prêts et épargne Données anthropométriques

9. Les questionnaires sont étroitement adaptés à chaque pays, ce qui explique les variations de contenu des instruments utilisés, ainsi que l'inclusion de nouveaux modules et de nouveaux thèmes au fil des années. Par exemple, en 2001 en Bosnie-Herzégovine, le module santé a été élargi avec l'inclusion de questions sur la dépression afin de mesurer l'incidence de cette maladie mentale et d'identifier les liens entre elle et certains aspects du bien-être et de la participation au marché du travail. Au Guatemala en 2000, un module sur le capital social a été ajouté pour recueillir des données sur les dimensions sociales de la pauvreté, notamment sur la participation aux programmes communautaires/gouvernementaux et aux actions collectives, sur les causes de l'exclusion sociale, sur les perceptions du bien-être et sur les perceptions de la justice et l'accès à la justice. En Afrique du Sud, en Albanie, au Brésil, au Népal, en Jamaïque et au Tadjikistan, des questions ont été ajoutées sur certaines mesures subjectives de la pauvreté, afin d'analyser le rapport entre ces mesures et d'autres⁶. Le tableau XXIII.2 présente certains modules qui ont été ajoutés ces dernières années. En résumé, si l'ensemble type de modules demeure, l'enquête LSMS effectuée dans chaque pays reflète les priorités du pays, les besoins en données et les préoccupations du pays au moment de l'enquête. Un récent projet de recherche de la Banque mondiale sur « L'amélioration de la pertinence des enquêtes LSMS sur les politiques » a conduit à la publication d'un nouvel ouvrage décrivant, par thème, les questions de politique générale auxquelles peuvent répondre les données des enquêtes LSMS et offrant des conseils pour la conception des questionnaires⁷.

⁶ Pour plus de renseignements sur le travail effectué sur le capital social au Guatemala, voir Banque mondiale (2002b). Pour plus de renseignements sur les mesures subjectives de la pauvreté, voir Pradhan et Ravallion (2000), Ravallion et Lokshin (2001), Ravallion et Lokshin (2002). L'analyse des données de la Bosnie-Herzégovine se poursuit.

⁷ Dans Grosh et Glewwe, eds. (2000).

10. Durant la phase de conception des questionnaires, il faut faire en sorte que les questions pertinentes soient identifiées et incorporées aux questionnaires. Dans la plupart des pays, un Groupe d'utilisateurs des données ou Comité de direction est formé avec des représentants des différents ministères d'exécution, des donateurs et des milieux universitaires, ainsi que de l'Office national de statistique (ONS). Ce groupe est chargé d'identifier les besoins en données pour la définition de certaines politiques, de manière à assurer la collecte de données pertinentes. La phase de conception des questionnaires dure en moyenne environ huit mois et fait intervenir le plus grand nombre d'acteurs possible. Ce processus assez laborieux offre l'avantage supplémentaire d'engendrer la demande de données et un sentiment de paternité à l'égard des données recueillies, ce qui favorise une plus grande utilisation de ces données dans l'élaboration des politiques.

Tableau XXIII.2

Exemples de modules additionnels

Thèmes	Pays et année
Activités de la vie quotidienne	Kosovo (2000), Kirghizistan (1993, 1996, 1997, 1998), Jamaïque (1995), Nicaragua (1993)
Incapacité	Nicaragua (1993)
Impact de la mortalité liée au sida	République-Unie de Tanzanie-Kagera (1991-1994)
Tests d'instruction élémentaire (alphabétisme et/ou arithmétique)	Viet Nam (1997-1998), Jamaïque (1990), Maroc (1990-1991)
Santé mentale	Bosnie-Herzégovine (2001)
Privatisation	Bosnie-Herzégovine (2001), Kirghizistan (1996, 1997)
Chocs/vulnérabilité	Bolivie (1999, 2000), Guatemala (2000), Paraguay (2000-2001), Pérou (1999)
Capital social	Guatemala (2000), Kosovo (2000), Panama (1997), Paraguay (2000-2001)
Mesures subjectives de la pauvreté	Albanie (2002), Brésil (1996), Jamaïque (1997), Népal (1996), Afrique du Sud (1993), Tadjikistan (1999)
Utilisation du temps	Guatemala (2000), Nicaragua (1998), Jamaïque (1993), Pakistan (1991), Maroc (1990-1991), République-Unie de Tanzanie-Kagera (1991-1994)

2. Questions relatives aux échantillons

11. En général, les enquêtes LSMS sont des enquêtes nationales qui utilisent des échantillons probabilistes à phases multiples de ménages⁸. Les échantillons d'ensemble sont petits (par rapport à ceux d'autres enquêtes), variant généralement de 2 000 à 5 000 ménages. Cette taille s'explique principalement par deux raisons. Premièrement, il y a le souci de qualité et la nécessité de doser entre l'erreur d'échantillonnage et l'erreur autre que celle-ci (voir l'analyse plus approfondie de cette question à la section C.4 ci-dessous). Deuxièmement, la visée analytique des enquêtes LSMS porte sur les déterminants ou les relations entre les caractéristiques des ménages et non pas sur des estimations précises de taux, de ratios ou de moyennes spécifiques. C'est pourquoi les échantillons LSMS restent raisonnablement modestes et, généralement, ne sont pas suffisamment grands pour que les résultats de l'enquête puissent être décomposés par région géographique telle qu'État, municipalité ou département.

12. Si l'échantillonnage probabiliste est utilisé dans toutes les enquêtes LSMS, la conception proprement dite varie selon le pays et la situation⁹. Les domaines de l'étude sont identifiés (urbain/rural, régions) et, à l'intérieur de chaque domaine, on utilise une conception en grappe stratifiée à deux phases¹⁰. Comme la plupart des enquêtes sur les ménages, les enquêtes LSMS utilisent un échantillon aléatoire simple [*simple random sample* (SRS)]. Cela tient à des considérations de coût, quoique les conceptions en grappes réduisent la précision des estimations (voir section E.4, ainsi que l'annexe III ci-après pour plus de renseignements sur les effets liés à la conception de l'échantillon qui découlent d'un échantillonnage à phases multiples). Les unités primaires d'échantillonnage (UPE) sont des unités définies géographiquement, sélectionnées selon une probabilité proportionnelle à leur taille. Le cadre d'échantillonnage est généralement le dernier recensement de la population du pays, quoique d'autres cadres aient été utilisés lorsque le recensement n'était pas disponible ou était sans rapport avec l'objet de l'enquête (voir, par exemple, les documents d'information de base pour l'Enquête LSMS de 1993 au Nicaragua, où les listes électorales complétaient les informations dépassées du recensement, et l'Enquête LSMS de 2001 en Bosnie-Herzég-

⁸ En fait, comme pour la plupart des enquêtes sur les ménages, c'est l'unité d'habitation qui est choisie et ce sont tous les ménages y habitant qui sont interrogés. Il convient de noter que lorsque l'on utilise une conception par panel, le choix entre l'unité d'habitation ou le ménage dépend de l'objet du panel et de questions de logistique.

⁹ Les Documents d'information de base de chaque enquête donnent des détails de la conception de l'échantillon utilisé pour l'enquête. Ces documents peuvent être consultés sur le site web LSMS : <http://www.worldbank.org/lms/>.

¹⁰ Toutefois, certains pays ont nécessité une conception en trois phases.

govine, qui a nécessité de vastes opérations d'établissement de listes, en raison de la guerre civile).

13. Une fois que les UPE ont été sélectionnées, on procède à leur énumération afin de s'assurer que l'on dispose d'une liste à jour et exacte de tous leurs logements et ménages. Cette liste est dressée aussi près que possible de la date à laquelle débute le travail de terrain nécessaire à l'enquête. Pour éviter tous risques de distorsions, elle est établie non pas par les enquêteurs eux-mêmes mais par le département de cartographie de l'ONS. Lorsque l'on dispose de la liste complète et à jour de tous les logements de l'UPE, on choisit systématiquement les unités secondaires d'échantillonnage (ménages), en général un nombre fixe de ménages au sein de chaque UPE, de l'ordre de 12 à 18. On collecte ensuite des données auprès de tous les membres de ces ménages. Bien que la conception de l'échantillon des enquêtes LSMS doive assurer une couverture nationale, dans certains cas, par suite d'un conflit civil ou d'une catastrophe naturelle, certaines zones peuvent en être exclues.

14. Généralement, les estimations des enquêtes LSMS nécessitent l'utilisation de pondérations d'échantillonnage. Même lorsque la conception originale de l'échantillon exige une forme d'autopondération, comme ce fut le cas au Ghana, au Nicaragua (1993) et au Tadjikistan, les différences de taux de non-réponse obligent à différencier les pondérations à utiliser dans l'analyse des données. En fait, la plupart des échantillons ne sont pas autopondérés. Souvent, la conception de l'échantillon dans un pays donné est fonction de considérations analytiques propres au pays. Par exemple, certains sous-groupes de population qui sont de taille modeste mais intéressent le gouvernement (minorités ethniques, régions reculées, personnes s'adonnant à une activité économique particulière ou se trouvant dans une zone couverte par un important projet du gouvernement) peuvent avoir à être suréchantillonnées de manière à assurer qu'il y ait suffisamment de cas pour permettre de les analyser séparément. Là encore, la conception de ces échantillons oblige à leur affecter des pondérations pour l'analyse des données. Enfin, étant donné la conception des échantillons utilisés dans les enquêtes LSMS, les tests statistiques importants effectués sur les données doivent aussi tenir compte du caractère à phases multiples de la conception (voir les chapitres de cette publication sur les effets de conception de l'échantillon pour plus de renseignements sur cette question).

3. Organisation sur le terrain

15. Comme on l'a vu précédemment, les objectifs des enquêtes LSMS conditionnent leur structure et leur teneur : ils se reflètent également dans les méthodes employées pour le travail de terrain. Ce travail est conçu de telle manière que les données sont recueillies par des équipes mobiles d'enquêteurs qui assurent également les activités d'entrée des données et une étroite supervision¹¹. Chaque ménage reçoit au moins deux visites à deux semaines d'intervalle l'une de l'autre. La figure XXIII.2 donne une illustration graphique de la façon dont s'effectue le travail de terrain. En un mois, chaque équipe d'enquêteurs effectue les entrevues dans deux communautés sélectionnées (UPE) : elle travaille dans la première communauté la première et la troisième semaine, et dans la seconde communauté la deuxième et la quatrième semaine. La première moitié du questionnaire est remplie lors de la première visite à chaque communauté. Entre les deux visites, les données de la première sont entrées et vérifiées pour y détecter les erreurs. La seconde visite est utilisée pour corriger les erreurs détectées à la suite de la première visite et procéder à la seconde moitié de l'enquête, et prévoit une période pour la collecte d'informations sur les dépenses alimentaires¹². Les données sont généralement recueillies sur une période totale de 12 mois, pour

¹¹ Pour plus de détails sur les équipes d'enquêteurs, voir annexe II.

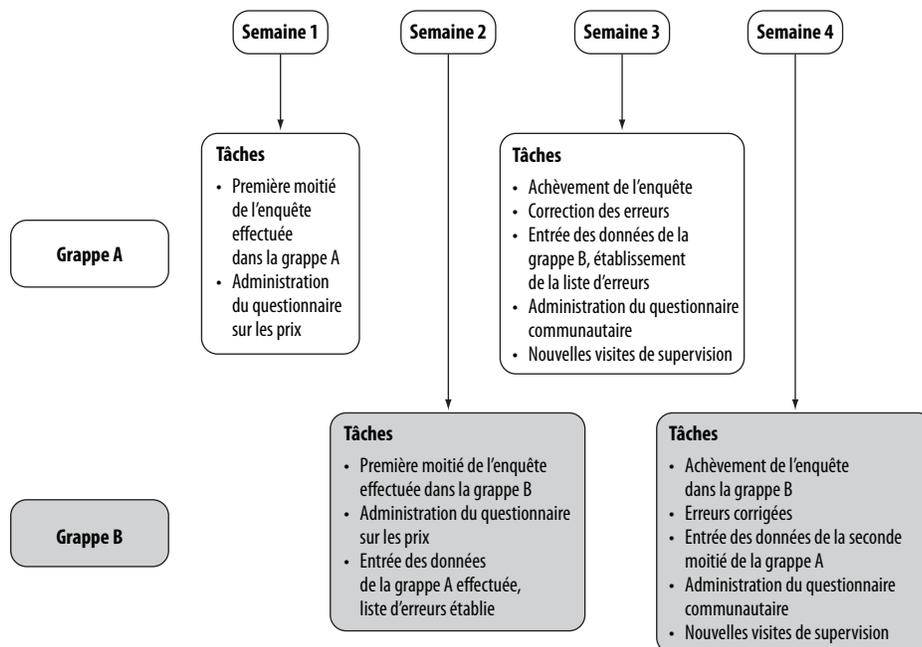
¹² Bien que deux visites soient officiellement prévues, l'utilisation de déclarants directs pour toutes les sections du questionnaire signifie qu'en fait les enquêteurs rendent visite à chaque ménage autant de fois qu'il le faut pour interroger tous les membres du ménage.

permettre, s'il y a lieu, les ajustements saisonniers, mais certains pays préfèrent des périodes plus courtes.

16. Le superviseur est responsable de l'administration des questionnaires communautaires et des questionnaires sur les prix, parallèlement à la collecte de données auprès des ménages de l'UPE par son équipe d'enquêteurs. Les enquêtes sur les services peuvent nécessiter un personnel supplémentaire.

Figure XXIII.2

Calendrier mensuel d'activités de chaque équipe



4. Qualité

17. L'une des préoccupations essentielles, lors de l'exécution d'enquêtes LSMS, est de veiller à l'excellente qualité des données recueillies. La complexité de l'enquête rend les mécanismes de contrôle de qualité d'autant plus importants. Comme le montre le tableau XXIII.3, le contrôle de qualité prend des formes diverses, des plus simples – fondées sur des questions mot pour mot, des questions à omettre explicitement, des questionnaires traduits dans les langues du pays, et des questions fermées pour limiter le plus possible les erreurs de l'enquêteur – à la plus complexe, qui consiste à entrer des données avec nouvelles visites immédiates aux ménages pour corriger les incohérences ou saisir les données manquantes. Il est clair que tous ces contrôles de qualité ne sont pas uniques aux enquêtes LSMS, mais étant donné la complexité de ces enquêtes, on s'est efforcé de prévoir un ensemble complet de contrôles de qualité. Outre les contrôles mentionnés précédemment, et peut-être de façon plus controversée, le programme LSMS a opté pour un échantillon de petite taille pour limiter les erreurs autres que les erreurs d'échantillonnage. La logique de cette démarche est que si les erreurs d'échantillonnage peuvent être grosses lorsque l'on utilise de petits échantillons, ces erreurs peuvent au moins être chiffrées. En revanche, les erreurs autres que les erreurs d'échantillonnage émanent de nombreuses sources, et leur ampleur est pratiquement impossible à mesurer; on sait pourtant que la totalité de ces erreurs tend à s'accroître à mesure qu'augmente la taille de l'échantillon. On a donc décidé de limiter ces erreurs, quitte à limiter le niveau de fragmentation géographique que permettent

les données de l'enquête. Du fait de l'importance accordée par les enquêtes LSMS à l'étude des relations entre divers aspects des niveaux de vie, de préférence à la mesure précise d'indicateurs ou de taux spécifiques, cette décision est moins gênante qu'elle ne pourrait l'être dans d'autres enquêtes¹³. Enfin, certaines méthodes récentes employées pour lier les données des enquêtes LSMS (et d'autres enquêtes) aux données du recensement qui permettent d'imputer la pauvreté à l'intérieur des données du recensement permettent d'atténuer quelque peu le problème posé par la petite taille de l'échantillon, du moins du point de vue des mesures de la pauvreté et de l'inégalité¹⁴.

¹³ Une enquête sur la main-d'œuvre, qui est censée afficher de très faibles variations des taux de chômage dans le temps, nécessite un échantillon beaucoup plus grand que celui nécessaire à l'analyse des déterminants du chômage, qui serait au centre de l'analyse de l'enquête LSMS.

¹⁴ Pour plus de détails sur cette technique, voir section E ci-dessous sur les utilisations des données des enquêtes LSMS.

Tableau XXIII.3

Contrôles de qualité dans les enquêtes LSMS

Domaines du contrôle de qualité	Contrôles
Questionnaire	Questions mot pour mot Enchaînement des questions Utilisation minimale de questions ouvertes Traduction écrite dans les langues pertinentes ^a Questions délicates placées à la fin Présentation : un formulaire pour toutes les données sur les ménages et les individus
Phase pilote	Test pilote formel du questionnaire et du travail de terrain
Déclarants directs	Individus et éléments les mieux informés
Entrée simultanée des données	Vérification de la portée, dépistage des incohérences Nouvelles visites aux ménages pour apporter des corrections
Format à deux temps	Réduit la fatigue Crée des périodes de rappel Permet de vérifier l'entrée des données et les corrections avec les ménages
Formation	Formation intensive des enquêteurs (un mois), des superviseurs et du personnel chargé d'entrer les données
Travail de terrain décentralisé	Équipes mobiles composées d'un superviseur, de deux ou trois enquêteurs, d'un préposé à l'entrée des données équipé d'un ordinateur et d'une imprimante et d'une voiture avec chauffeur
Supervision	Un superviseur pour deux ou trois enquêteurs
Échantillon de petite taille	Limite les erreurs autres que les erreurs d'échantillonnage
Politique d'accès aux données	Ouvert à tous les chercheurs et institutions

^a Dans les pays où il n'existe pas de formulaires dans certaines langues (langues vernaculaires au Panama, par exemple), on utilise des enquêteurs bilingues. Cette solution n'est pas idéale et est à éviter, à moins qu'elle ne soit absolument nécessaire.

¹⁵ Même pour les répondants les « mieux informés », la durée effective de l'entrevue est maintenue à moins d'une heure, jugée la durée maximum pendant laquelle une personne devrait être interrogée. Toutefois, pour certains ménages, cette durée peut être dépassée, et il faut éviter de fatiguer le déclarant et veiller à ce que la qualité des données ne se dégrade pas avec la durée.

¹⁶ Dans le cas des enfants de moins de 10 ans ou âgés de 12 ans, ou des membres de la famille incapables de communiquer, on peut utiliser des substituts. Dans ce cas, il faut noter le code d'identification du vrai répondant.

18. Le recours à des déclarants directs, également appelés autorépondants, est un autre mécanisme de contrôle de qualité des enquêtes LSMS, qui présente deux avantages essentiels : il allège la charge imposée à tout répondant donné et réduit donc sa fatigue. Le questionnaire pour les ménages est en fait une série de brèves entrevues individuelles (de 10 à 15 minutes), les répondants les mieux informés sur la consommation, l'agriculture et le commerce familial étant les seuls à être soumis à des entrevues plus longues¹⁵. L'utilisation de déclarants directs permet aussi d'améliorer la qualité des données recueillies en faisant en sorte que seule la personne la mieux informée réponde aux questions posées¹⁶. On ne peut compter qu'un membre quelconque du ménage puisse donner des réponses exactes et complètes à la fois sur la santé, l'éducation, le travail, la migration, le crédit et la fécondité ou sur toutes les activités des autres membres de la famille, ce serait trop demander. De plus, au sein d'une même famille, certains peuvent avoir intérêt à ne pas dévoiler certaines informations aux

autres membres (le crédit, l'épargne, les revenus et l'utilisation de contraceptifs sont autant de domaines sur lesquels certains peuvent juger préférable de ne rien dévoiler au reste de la famille). Le recours à des déclarants directs est donc le seul moyen d'assurer une information exacte de la part de chaque membre de la famille. Dans la mesure du possible, les enquêteurs apprennent à procéder à des entrevues individuelles sans témoin.

19. La formation de tout le personnel participant à chaque enquête LSMS est un autre mécanisme de contrôle de qualité. Cette formation est dispensée « sur le tas » au personnel de l'ONS et dans le cadre de cours plus formels en tant que de besoin. Pour le personnel de terrain, les enquêteurs, les superviseurs et les préposés à l'entrée des données, des ressources substantielles sont investies dans la formation formelle. En règle générale, la formation de ce personnel dure quatre semaines et comprend des cours théoriques et des travaux pratiques. Lorsqu'elle est achevée, les membres du personnel de terrain sont sélectionnés parmi ceux qui ont suivi avec succès cette formation. Les résultats sont généralement jugés satisfaisants à la fois sur la base d'une participation fructueuse aux cours et du succès à un examen final.

20. Un dernier moyen d'améliorer la qualité des données auquel il est fréquent que l'on ne songe pas consiste à promouvoir le libre accès aux microdonnées résultant de l'enquête. Le fait d'assurer une ample utilisation des ensembles de données par un large éventail de chercheurs et de décideurs permet de vérifier soigneusement les données recueillies et, en créant des boucles de rétroaction avec les sources de données, d'améliorer la qualité des enquêtes futures. Des accords de libre accès ont été conclus pour la plupart des ensembles de données des enquêtes LSMS, et l'on aide les gouvernements à diffuser ces données. Bien que la Banque mondiale n'ait pas la propriété de ces ensembles de données, elle a reçu l'autorisation de diffuser directement plus de la moitié d'entre eux (en fait, 30% des ensembles de données peuvent être téléchargés directement du site web LSMS)¹⁷. Quant aux autres, la plupart d'entre eux peuvent être diffusés une fois que le Gouvernement approuve chaque demande individuelle. D'après ceux qui ont demandé une telle autorisation, celle-ci est accordée dans environ 90 % des cas.

¹⁷ <http://www.worldbank.org/lsms/>.

5. Entrée des données

21. L'entrée immédiate des données nécessite l'utilisation d'un logiciel perfectionné d'entrée des données qui vérifie les erreurs d'ordre de grandeur, les contradictions à l'intérieur d'un même ensemble de données et entre différents ensembles et, dans la mesure du possible, par rapport à des tableaux de référence externes (par exemple, tableaux donnant des données anthropométriques ou des données sur les rendements et les prix de certaines récoltes). Les données sont entrées sur le terrain, à l'aide d'ordinateurs portables, durant leur phase de collecte, et les préposés à l'entrée des données font partie intégrante des équipes mobiles chargées de l'enquête. Elles sont entrées immédiatement après chaque entrevue, et une liste d'erreurs, d'incohérences et d'informations manquantes est établie à partir des données ainsi entrées. L'enquêteur retourne alors auprès de la famille pour clarifier tout problème et combler les lacunes constatées. Cette méthode permet d'éviter un laborieux nettoyage des données après l'enquête. Il est préférable d'éviter ce nettoyage : s'il tend à produire des ensembles de données cohérents, ces ensembles ne sont pas ceux qui donnent l'image la plus fidèle de la situation de chaque individu. Il demande en outre beaucoup de temps, et retarde d'autant l'utilisation des données et, au pire, rend certaines de celles-ci dépassées. Avec l'apparition d'ordinateurs peu coûteux et puissants et de nouveaux logiciels, il est probable que certaines enquêtes LSMS s'effectueront de manière entiè-

rement électronique à partir de méthodes d'entrevues individuelles assistées par ordinateur [*computer-assisted personal interview (CAPI) methods*]. Telle est la voie qui est actuellement étudiée, vu la possibilité qu'elle offre de réduire le délai entre le travail de terrain et la publication des données, et d'améliorer la qualité de celles-ci¹⁸.

6. Durabilité

22. Dans la version la plus simple des enquêtes LSMS, les trois plus gros obstacles à leur durabilité et à leur utilisation à long terme ainsi qu'à celle des données en résultant pour l'élaboration des politiques sont les contraintes budgétaires, la rotation du personnel et le manque de capacité d'analyse. Si l'on ne dispose pas de remède miracle pour assurer la durabilité des enquêtes LSMS, l'expérience offre plusieurs indications de la façon d'accroître les possibilités d'y parvenir. Les premières de ces indications font ressortir la nécessité d'associer les décideurs et les utilisateurs des données aux phases de conception et d'analyse. Tel est essentiellement le point de départ du processus de création d'une demande à l'égard des résultats des enquêtes LSMS et de l'utilisation des données de ces enquêtes dans les décisions de politique générale. Comme ce sont ces usagers finals qui bénéficient de ces données (et non pas l'ONS lui-même), ce sont eux qui ont le plus intérêt à ce qu'il soit répondu aux besoins de ressources budgétaires pour les enquêtes à venir lors du processus d'établissement du budget de l'État. Souvent, la durabilité dépend de l'apparition d'un ou de plusieurs « champions » de la cause de l'enquête, en dehors du système statistique proprement dit¹⁹.

23. La deuxième grande leçon est que la durabilité ne s'obtient qu'au terme d'un long processus : l'investissement dans des enquêtes ponctuelles n'a qu'un faible impact à long terme. Il faut une action plus systématique prolongée sur plusieurs années pour former une masse critique d'agents, démontrer l'efficacité et l'utilisation des instruments, créer des liens entre producteurs et utilisateurs et adapter la méthodologie aux besoins et aptitudes du pays. Par ailleurs, en investissant dans une documentation appropriée des enquêtes et dans l'archivage et la diffusion des données, on protège la mémoire institutionnelle des effets des départs de certains membres du personnel. Près de 40 % des pays qui ont effectué une enquête LSMS en ont effectué d'autres.

24. Enfin, pour constituer une capacité d'analyse, il faut définir un objectif explicite²⁰. Cela permet d'accroître l'utilisation des données et facilite la création d'une demande de futurs ensembles de données. De plus, le développement des compétences du personnel de l'ONS et, par là même, le renforcement du prestige de l'Office au sein du gouvernement encouragent le personnel à la stabilité²¹. Par ailleurs, certaines forces extérieures peuvent aussi contribuer à accroître la demande de données. Les stratégies de réduction de la pauvreté conçues par les pays qui reçoivent des prêts concessionnels de la Banque mondiale et du Fonds monétaire international (FMI), et les Objectifs du Millénaire pour le développement nécessitent tous des données sur la mesure et le suivi de la pauvreté et des principaux indicateurs sociaux. Le fait qu'il s'agisse d'objectifs à long terme peut favoriser l'apparition de systèmes de suivi et d'évaluation fortement tributaires des enquêtes sur les ménages telles que les enquêtes LSMS, en même temps que de données administratives et de données de projets²². L'évaluation récente par la Banque interaméricaine de développement, la Banque mondiale et la Commission économique pour l'Amérique latine et les Caraïbes (CEPALC) d'un projet visant à améliorer les enquêtes sur les ménages²³ fait ressortir le caractère à long terme de la durabilité et pose la question supplémentaire du passage du financement par les donateurs au financement local qui doit également retenir l'attention²⁴.

¹⁸ L'utilisation des méthodes CAPI est l'un des facteurs qui contribuent à l'aptitude du United States Bureau of the Census à publier les résultats de son enquête mensuelle sur la main-d'œuvre (*Current Population Survey*) dans les 10 jours qui suivent le travail de terrain. Une expérience visant à comparer les coûts et les avantages des méthodes CAPI à ceux de l'entrée immédiate des données des enquêtes LSMS est envisagée pour l'Albanie en 2003.

¹⁹ La Jamaïque offre un exemple de cette approche. À l'origine, la demande émanait du cabinet du Premier Ministre, et le Ministère du plan a participé à chaque phase de la conception et de l'exécution de l'enquête avec l'Office de statistique comme agent d'exécution. L'enquête LSMS a lieu tous les ans depuis la fin des années 1980 en Jamaïque. Voir Grosh (1991) pour plus de détails sur cet exemple.

²⁰ Un résumé des leçons tirées des enquêtes LSMS en ce qui concerne la constitution d'une capacité d'analyse est présenté par Blank et Grosh (1999).

²¹ Il y a toujours le souci de maintenir la séparation entre les collecteurs et les analystes des données. Il ne faut pas perdre de vue les questions de crédibilité lorsque l'on fait preuve de laxisme à l'égard de cette séparation.

²² La création du Partenariat statistique au service du développement à l'aube du XXI^e siècle (PARIS 21), qui vise à favoriser l'amélioration des données, souligne l'importance de la collecte, de l'analyse et de l'utilisation de données durables.

²³ Le projet Banque interaméricaine de développement-Banque mondiale-CEPALC est intitulé « Amélioration des enquêtes sur les conditions de vie », mais est connu plus communément sous son sigle espagnol MECOVI.

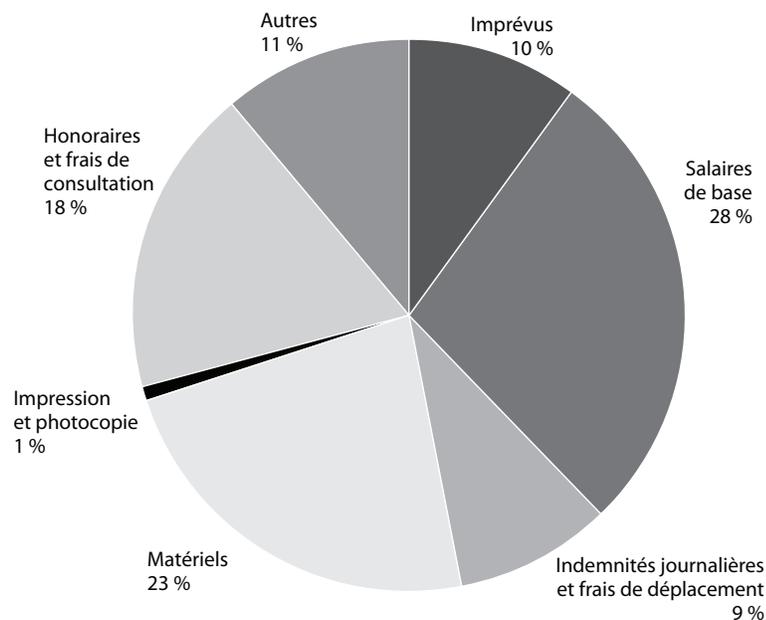
²⁴ Voir Ryten (2000).

D. Coûts d'exécution d'une enquête LSMS

25. L'attention portée à la qualité a de sérieuses répercussions sur les coûts des enquêtes en temps et en ressources. Les coûts des enquêtes LSMS varient de 400 000 à 1,5 million de dollars, selon le pays et l'année. Par ménage, cela est comparable à d'autres enquêtes complexes telles que les Enquêtes sur les revenus et les dépenses et les Enquêtes démographiques et sanitaires. Bien entendu, ces coûts varient en fonction de la capacité de l'ONS, de l'état de l'infrastructure statistique en place, des objectifs de l'enquête et de la difficulté des déplacements à l'intérieur du pays. Ils sont sensiblement plus faibles lorsque l'agence d'exécution dispose déjà d'une bonne infrastructure et d'un personnel expérimenté. Généralement, les crédits alloués à chaque enquête proviennent de sources très diverses : du budget de l'État (pour l'ONS et d'autres institutions), ou de dons et de crédits bilatéraux et multilatéraux. Parfois, le secteur privé finance également une partie des coûts de l'enquête²⁵.

26. En général, le coût d'une enquête LSMS est fonction de la méthode employée, de la taille de l'échantillon et de la complexité du travail de terrain. La figure XXIII.3 indique les éléments du coût d'une enquête LSMS et la pondération relative de chacun de ces éléments²⁶. (L'annexe II propose un simple exercice destiné à aider le lecteur à élaborer un budget pour une enquête LSMS.)

Figure XXIII.3
Éléments de coût d'une enquête LSMS (part du coût total)



Source : D'après Grosh et Muñoz (1996), tableau 8.2.

27. Le plus gros élément de coût est constitué par les salaires. Près des trois quarts de cet élément correspondent aux salaires du personnel de terrain : enquêteurs, superviseurs, préposés à l'entrée des données, anthropométristes et chauffeurs. Le personnel de terrain nécessaire aux enquêtes LSMS est nombreux (par rapport à la taille de l'échantillon) en raison du ratio élevé superviseur/enquêteurs (généralement de 1 pour 3), de la longueur du questionnaire et du recours à des déclarants directs qui limite le nombre de ménages qui peuvent être visités chaque jour, de l'inclusion de l'en-

²⁵ Par exemple, au Pérou, un espace limité du questionnaire est réservé à des entreprises privées ou des chercheurs qui paient pour que soient ajoutées certaines questions particulières au questionnaire d'un trimestre donné.

²⁶ La présente section sur les coûts s'appuie sur Grosh et Muñoz (1996) et sur un exposé présenté par Juan Muñoz à un cours de la Banque mondiale sur la pauvreté et l'inégalité, 26-28 février 2002.

trée des données dans les travaux des équipes de terrain et de la fourniture de moyens de transport à chaque membre de l'équipe pour assurer la mobilité et l'intégrité de l'équipe. Les autres salaires sont ceux du personnel de bureau : il s'agit généralement de personnel de l'ONS, encore qu'un coordonnateur du projet puisse être engagé de l'extérieur en cas de besoin.

28. En second viennent les coûts de matériel et d'équipement. Cela couvre les ordinateurs et les véhicules (achetés ou loués), et l'entretien, ainsi que les autres matériels de bureau. Cette composante est celle qui varie le plus largement en fonction de l'infrastructure dont dispose l'ONS ou l'agent d'exécution. Par ailleurs, les sources de financement peuvent alourdir les coûts si elles interdisent l'achat de véhicules : la location des véhicules nécessaires peut parfois s'avérer sensiblement plus coûteuse.

29. L'assistance technique est le troisième élément de coût. Celui-ci varie lui aussi très sensiblement selon les compétences et l'expérience de l'agence d'exécution. Bien entendu, les pays qui en sont à leur deuxième ou troisième enquête LSMS ont beaucoup moins besoin d'assistance technique et d'équipement. En général, les types d'activités nécessitant le plus d'assistance technique sont l'échantillonnage, la conception de questionnaires, l'adaptation de l'entrée des données au contexte, l'organisation du travail de terrain et les techniques d'analyse.

30. Bien entendu, les coûts de l'enquête LSMS sont justifiés si cette enquête produit vraiment des données de meilleure qualité qui peuvent servir à améliorer la politique suivie. Bien que coûteuses en valeur absolue, en regard de l'ampleur des dépenses consacrées à la politique sociale, ces enquêtes peuvent se révéler très économiques. La section suivante en donne la preuve en examinant la qualité des données de récentes enquêtes LSMS. La qualité est définie par rapport aux données manquantes, à l'utilité des données aux fins de l'enquête LSMS, à l'homogénéité interne et aux effets de conception.

E. Quel est le degré d'efficacité de la conception des enquêtes LSMS sur leur qualité ?

1. Taux de réponse

31. Le premier critère de qualité de l'enquête est le taux global de réponse : les ménages sélectionnés pour l'échantillon répondent-ils à l'enquête ou un nombre élevé d'entre eux s'abstiennent-ils, en faussant ainsi le résultat final²⁷ ? Il est utile d'examiner les taux de réponse, car ils donnent une indication de la valeur de la formation, de la bonne conception du questionnaire et de la compétence des enquêteurs, ainsi que de la qualité des procédures de sélection de l'échantillon (énumération, tenue à jour des cartes, etc.). Les enquêtes LSMS de nombreux pays ont donné des taux de réponse remarquablement élevés. Le tableau XXIII.4 indique les taux de réponse d'enquêtes récentes. Cela dit, les enquêtes LSMS ne sont pas à l'abri des effets de certaines situations propres au pays. Dans les pays sortant d'un conflit, où les niveaux de confiance à attendre sont faibles, les taux de réponse sont plus faibles, comme en témoignent les enquêtes LSMS effectuées en Bosnie-Herzégovine et au Kosovo. Cependant, ce sont peut-être les faibles taux de réponse observés en Jamaïque qui illustrent le mieux l'importance de la qualité des mécanismes de contrôle. La Jamaïque utilise moins les techniques courantes de terrain et les mesures de contrôle de qualité employées pour les enquêtes LSMS : cela ne semble pas s'être traduit par des taux de réponse plus faibles²⁸. Au Guatemala, le faible taux de réponse est dû probablement au long délai qui s'est écoulé entre la date d'achèvement des listes de ménages et les dates auxquelles

²⁷ Les taux de participation correspondent aux ménages dans leur intégralité et non pas à certains de leurs membres.

²⁸ Voir Banque mondiale (2001) pour plus de détails sur le travail effectué sur le terrain pour l'Enquête sur les conditions de vie en Jamaïque.

l'enquête s'est déroulée sur le terrain. Pour les dernières entrevues, ce délai a été de neuf à dix mois.

Tableau XXIII.4

Taux de réponse enregistrés lors de récentes enquêtes LSMS

Pays	Année	Nombre de logements sélectionnés	Taille effective de l'échantillon	Taux de réponse ^a (en pourcentage)
Bosnie-Herzégovine	2001	5 400	5 402	82,6
Ghana ^b	1998-1999	6 000	5 998	97,4
Guatemala	2000	8 940	7 468	83,5
Jamaïque	1999	2 540	1 879	74,0
Kosovo	2000	2 880	2 880	82,0
Kirghizistan	1998	2 987	2 979	99,7
Nicaragua	1998-1999	4 370	4 209	96,3
Tadjikistan	1999	2 000	2 000	..
Viet Nam	1997/98	5 994	5 999	93,9

^a La Bosnie-Herzégovine, le Ghana, le Kosovo, le Tadjikistan et le Viet Nam ont remplacé les ménages défaillants. Le taux de réponse est fondé sur le nombre d'entrevues menées à leur terme, moins les ménages utilisés en remplacement, divisé par la taille prévue de l'échantillon. En Bosnie-Herzégovine, 938 ménages de remplacement ont été utilisés, contre 155 au Ghana, 519 au Kosovo et 372 au Viet Nam. Les auteurs n'ont pas réussi à déterminer le nombre de ménages de remplacement utilisés au Tadjikistan.

^b L'enquête effectuée au Ghana a été menée sur la base de sept visites rendues à chaque ménage. Le chiffre donnant la taille finale de l'échantillon est celui des ménages qui ont effectivement reçu sept visites.

Note : Deux points (..) indiquent l'absence de données.

2. Non-réponse ponctuelle

32. Le pourcentage de non-réponse ponctuelle est un autre indicateur de qualité. L'examen de cette question sur les trois premières enquêtes LSMS a révélé que ce taux est relativement insignifiant (moins de 1 % des réponses faisaient défaut pour 10 variables clés)²⁹. Il est également intéressant de comparer les taux de non-réponse ponctuelle enregistrés dans les enquêtes LSMS à ceux enregistrés dans d'autres enquêtes qui ne disposent pas des mêmes mécanismes de contrôle de qualité. Bien que cela ne soit pas toujours possible, on trouvera ici une petite comparaison. Un examen effectué en 1998 des enquêtes sur la main-d'œuvre [*Labour Force Surveys* (LFS)] en Amérique latine a permis de réunir des informations sur la fréquence des valeurs manquantes concernant le revenu des travailleurs salariés, des travailleurs indépendants et des employeurs³⁰. Trois des pays cités ont également effectué des enquêtes LSMS dans l'année qui a suivi leurs enquêtes sur la main-d'œuvre. Comme le montre le tableau XXIII.5, dans ces pays, les enquêtes LSMS ont fait sensiblement mieux, ou du moins aussi bien, que les enquêtes sur la main-d'œuvre sur la plupart des points de comparaison. Bien qu'il ne s'agisse que d'un exemple limité, cela semble montrer l'effet positif de l'investissement des enquêtes LSMS dans le contrôle de qualité.

33. Au lieu d'indiquer simplement le nombre de réponses manquantes, on pourrait peut-être mieux mesurer la qualité des données à la mesure dans laquelle elles peuvent être utilisées. Pour les enquêtes LSMS, qui ont pour principal but de mesurer le bien-être, il est particulièrement utile de déterminer dans quelle mesure les données recueillies conviennent à l'objectif visé. La mesure monétaire du bien-être qui est le plus couramment utilisée, pour ses avantages théoriques et pratiques, est la consommation

²⁹ Dans Grosh et Glewwe (1995).

³⁰ Dans Feres (1998). Bien que le revenu ne soit pas l'objet des enquêtes sur la main d'œuvre ou des enquêtes LSMS, les informations sur le revenu sont collectées de façon similaire dans ces deux types d'enquêtes.

totale des ménages. C'est une mesure complexe qui exige des données d'un ensemble de modules sur le questionnaire, tant au niveau des individus qu'à celui des ménages. Le plus souvent, les données sur la consommation sont tirées du module « logement » (utilisation de la valeur du logement, des services d'utilité publique et des autres dépenses de logement), du module « biens de consommation durable » (pour calculer la valeur des services utilisés, du module « éducation » (écoles privées, argent de poche), du module « consommation alimentaire » (aliments achetés, produits par le ménage ou reçus en cadeau), du module « agricole » (pour les aliments produits et consommés par la famille s'ils n'ont pas été saisis par le module « consommation alimentaire » et des modules « dépenses non alimentaires » (du savon aux articles de maison).

Tableau XXIII.5

Fréquence des données manquantes dans les enquêtes LSMS et les enquêtes LFS

Pays	Enquête	Pourcentage de données manquantes sur le revenu des			Pourcentage des déclarants directs
		Travailleurs salariés	Travailleurs indépendants	Employeurs	
Équateur	LFS, 1997	6,3	6,7	13,2	..
	LSMS, 1998	3,6	8,5	6,5	96,5
Nicaragua	Urban LFS, 1997	1,0	1,4	5,7	..
	LSMS, 1998	1,1	1,0	4,7	84,6
Panama	LFS 1997	2,9	36,2	26,0	..
	LSMS, 1996	1,0	3,5	8,4	98,7

Source : Les informations sur les Enquêtes sur la main-d'œuvre (LFS) émanent de Feres (1998); pour les enquêtes LSMS, les calculs ont été effectués par les auteurs.

Note : Pour le Nicaragua, dans l'enquête LSMS de 1998, le pourcentage de données manquantes n'inclut pas les zéros, car les instructions données aux enquêteurs demandaient à ceux-ci de donner zéro comme réponse si la personne avait perçu un revenu en nature et non en espèces. Cette catégorie englobait les agriculteurs, dont le revenu était calculé ailleurs dans le module sur la production agricole.

Note : Deux points (..) indiquent l'absence de données.

34. Le tableau XXIII.6 indique le pourcentage de familles pour lesquelles il a été possible d'établir un tel agrégat de consommation. Pour la plupart des enquêtes, très peu de ménages ont dû être éliminés de l'analyse, faute de données. L'exception est le Ghana. On ne sait pas exactement quel a été le problème dans le cas de l'enquête de 1998 au Ghana : l'échantillon était légèrement plus grand que dans les autres pays mais pas aussi nettement plus grand que dans le cas du Guatemala. Le fait que certaines données sur la consommation alimentaire aient été recueillies par la tenue d'un journal (et non pas selon la méthodologie habituelle aux enquêtes LSMS) a pu y être pour quelque chose : malheureusement, la documentation relative à l'enquête ne dit rien à ce sujet³¹.

³¹ Voir Ghana Statistical Service (2000).

3. Vérification de l'homogénéité interne

35. Il est d'une importance cruciale d'assurer l'homogénéité interne des données. C'est parce que la complexité des instruments de l'enquête ne permet pas de surveiller facilement cette homogénéité pendant les entrevues que tant de contrôles de qualité traitent de ces questions. Le tableau XXIII.7 présente trois exemples de vérifications de cette homogénéité. La première vérification détermine à quel point le questionnaire communautaire peut être lié aux données sur les ménages. La deuxième mesure le pourcentage des enfants d'âge préscolaire ou scolaire, tels qu'ils figurent sur la liste, qui disposent d'informations complètes sur leur scolarisation. Et la troisième dé-

termine si les personnes identifiées comme travailleurs indépendants dans le module sur la main-d'œuvre ont fourni des renseignements sur leurs activités dans le module « activités non agricoles du ménage ».

Tableau XXIII.6
Ménages avec agrégat complet de consommation :
exemples tirés de récentes enquêtes LSMS

Pays	Année	Taille de l'échantillon final	Ménages avec agrégat complet de consommation
			(pourcentage)
Bosnie-Herzégovine	2001	5 402	99,9
Ghana	1998-1999	5 998	87,7
Guatemala	2000	7 468	97,4
Jamaïque	1999	1 879	99,8
Kosovo	2000	2 880	100,0
Kirghizistan	1998	2 979	99,4
Nicaragua	1998-1999	4 209	96,0
Tadjikistan	1999	2 000	100,0
Viet Nam	1997-1998	5 999	100,0

Tableau XXIII.7
Homogénéité interne des données : liens réussis entre les modules
(en pourcentage)^a

Pays	Enquête sur les ménages et enquête communautaire ^a	Lien correct entre:		
		Liste et module « éducation » ^b		Module « emploi » et module « activités non agricoles du ménage » ^c
		Pré-scolaire	Primaire	
Bosnie-Herzégovine	...	99,5	99,8	90,4
Ghana	99,9	..	96,5	70,2
Guatemala	100	100	100	93,0
Jamaïque	96,4	..
Kosovo	100	..	100	58,6
Kirghizistan	100	86,5	98,4	93,1
Nicaragua	..	97,9	97,5	62,0
Tadjikistan	100	..	99,9	..
Viet Nam	100	..	99,6	98,1

Note : Deux points (..) indiquent l'absence de données.

^a Comparaison des ménages avec les communautés dans lesquelles ils vivent.

^b Comparaison de la variable âge sur la liste avec les individus présents dans le module « éducation ».

^c Comparaison des personnes ayant indiqué qu'elles étaient des travailleurs indépendants dans le module « emploi » avec la présence d'informations dans le module « activités non agricoles des ménages ».

Notes : Ce tableau se réfère au pourcentage de liens corrects. La Bosnie-Herzégovine, la Jamaïque et le Tadjikistan n'avaient pas de questionnaires communautaires. La Jamaïque, le Kosovo, le Tadjikistan et le Viet Nam n'avaient pas de module spécial sur les enfants d'âge préscolaire. La Jamaïque et le Tadjikistan n'ont pas recueilli d'informations sur les activités non agricoles des ménages.

36. Comme le montre le tableau, les deux premières vérifications indiquent une excellente qualité des données. En revanche, la troisième laisse apparaître quelques difficultés, dues à l'absence de contrôles appropriés sur le terrain entre les deux visites aux ménages. Le Viet Nam est le seul à avoir prévu une question précise à poser par l'enquêteur lors de la seconde visite pour s'assurer qu'il serait répondu aux ques-

tions de ce module. Il est clair qu'une telle mesure de précaution s'impose pour toutes les enquêtes.

4. Effets liés à la conception de l'échantillon

37. Le dernier critère d'évaluation des enquêtes LSMS a trait à la taille et à la conception de l'échantillon. Lorsque l'on utilise les données d'une enquête sur les ménages fondée sur un échantillon de conception complexe à phases multiples, stratification et mise en grappes, on calcule la vraie variance des estimations en tenant compte de ces caractéristiques de l'échantillon ainsi que de la pondération. L'effet de conception est le ratio de la vraie variance d'une estimation, compte tenu du caractère à phases multiples de l'échantillon, à la variance de l'estimation qui aurait été obtenue si l'on avait utilisé un échantillon aléatoire simple de la même taille³². Autrement dit, un effet de conception de 1 indique qu'il n'y a pas eu de perte de précision dans les estimations de l'échantillon du fait de son caractère d'échantillon à phases multiples, tandis qu'un effet de conception supérieur à 1 indique une baisse de l'efficacité de l'échantillon et de la précision des estimations avec l'utilisation d'un échantillon à phases multiples.

38. Dans le cadre des activités des enquêtes LSMS, on a procédé à une évaluation des effets de conception sur les principales variables et les principaux indicateurs sur certaines des premières de ces enquêtes. Cette évaluation, effectuée par Temesgen et Morganstein (2000), a fait ressortir plusieurs points essentiels dont il faut tenir compte quand on utilise les données d'une enquête LSMS (et, bien entendu, les données d'autres enquêtes sur les ménages utilisant des échantillons à phases multiples) et quand on conçoit les échantillons³³. Le premier point est que la multiplicité des thèmes des enquêtes LSMS complique le processus de conception de l'échantillon. Les effets de conception varient très largement, aussi bien entre les variables au niveau des individus qu'entre les variables au niveau des ménages, comme le montre le tableau XXIII.8, extrait des travaux de Temesgen et Morganstein (2000). En bref, le fait de réduire l'effet de conception d'une variable peut très bien accroître celui d'autres variables. Deuxièmement, le choix entre erreurs d'échantillonnage et autres erreurs est clair. Les effets de conception peuvent être profonds dans les enquêtes LSMS. Ce tableau montre que, dans la mesure où ces enquêtes sont utilisées pour établir des moyennes, des ratios et des estimations ponctuelles, il est capital de tenir compte de la conception de l'échantillon et de veiller avec soin à ce que les données soient bien utilisées.

Tableau XXIII.8
Exemples d'effets de conception dans les enquêtes LSMS

Pays	Consommation par habitant			Accès aux soins de santé			Taux de chômage		
	Pop. totale	Rurale	Urbaine	Pop. totale	Rurale	Urbaine	Pop. totale	Rurale	Urbaine
Côte d'Ivoire, 1988	6,7	3,6	5,5	6,3	5,7	2,2	7,0	4,4	5,7
Ghana, 1987	1,9	3,1	1,8	2,9	3,0	5,0	1,7	1,5	2,0
Ghana, 1988	3,2	2,9	2,9	2,2	2,5	3,6	1,3	1,1	1,4
Pakistan, 1991	1,6	1,1	2,6	5,0	4,0	5,2	4,6	4,7	2,5

Source : Temesgen et Morganstein (2000).

39. Comme avec les autres enquêtes, il est intéressant de noter que l'effet de conception varie non seulement selon les variables mais aussi géographiquement à l'intérieur d'un même pays pour une même variable et dans le temps pour une variable donnée. Enfin, les effets de conception peuvent présenter d'énormes différen-

³² Pour tous renseignements supplémentaires sur les questions de conception de l'échantillon, voir annexe III du présent chapitre et autres chapitres.

³³ L'annexe III du présent chapitre contient plusieurs des tableaux du rapport de Temesgen et Morganstein.

ces d'un pays à l'autre. Un examen attentif des corrélations intra-classe et des effets de conception observés lors d'enquêtes précédentes, le cas échéant, devrait permettre d'améliorer la conception des enquêtes LSMS à venir. Il faut faire preuve de prudence en présentant et en interprétant les résultats des enquêtes LSMS et d'autres enquêtes utilisant des échantillons à phases multiples, car la conception de l'échantillon utilisé peut être compliquée.

F. Utilisations des données d'enquêtes LSMS

40. Depuis des années, les données des enquêtes LSMS sont utilisées pour tout un ensemble de besoins de politique générale et de recherche. Certaines de ces applications sont commentées ailleurs³⁴, et le lecteur intéressé trouvera une bibliographie riche, quoique partielle, d'études et de rapports fondés sur ces données sur le site web LSMS. Cette bibliographie illustre les vastes possibilités d'utilisation de ces données aux fins d'analyse, et ces possibilités ne se limitent pas à celles indiquées. Les travaux constants de recherche et les révisions et amendements apportés aux questionnaires indiquent à quel point les applications ne cessent de changer. Pour illustrer la diversité des façons dont les données des enquêtes LSMS sont utilisées et combinées à d'autres données, il serait peut-être plus opportun de considérer une utilisation particulière — par exemple, le ciblage des programmes gouvernementaux sur les pauvres — que de tenter un examen exhaustif des utilisations de ces données.

41. Tout d'abord, un premier exemple tiré de la Jamaïque montre comment une analyse simple peut fournir à l'État de claires indications des effets du ciblage des pauvres en considérant d'autres programmes. Dans le cas de la Jamaïque, décrit dans Grosh (1991), il existait trois grands programmes de nutrition : la distribution générale de subventions alimentaires et de bons d'alimentation et les programmes de distribution de repas dans les écoles. L'enquête LSMS effectuée en Jamaïque a permis de chiffrer la valeur des avantages procurés aux familles pauvres par ces trois programmes et a montré que les subventions alimentaires, à la différence des deux autres programmes, était fortement régressive. Cette analyse a été l'une des raisons de la décision d'éliminer ces subventions et d'accroître les ressources affectées aux deux autres programmes.

42. Le deuxième outil que permet de créer les données des enquêtes LSMS sert au ciblage géographique sur les régions pauvres. En s'appuyant sur les données du recensement, on peut utiliser les données de l'enquête LSMS pour dresser des cartes de la pauvreté afin d'orienter les ressources et les programmes vers les régions pauvres³⁵. Pour que cette méthode puisse être utilisée, il faut que l'enquête LSMS et le recensement ne soient effectués qu'à quelques années d'écart³⁶. L'enquête LSMS donne une mesure fiable du bien-être (consommation totale des ménages); toutefois, en raison de la petite taille de l'échantillon, l'aptitude à fractionner les données sur la pauvreté se limite à la différenciation entre zones urbaines et rurales et entre certaines grandes régions géographiques du pays. Il est clair que cela n'est pas suffisant pour les gouvernements qui cherchent à orienter les ressources vers les poches de pauvreté et que, dans les systèmes décentralisés, cela ne facilite pas l'allocation des ressources aux pouvoirs locaux. De plus, les grandes régions se caractérisent souvent par une grande hétérogénéité, quant aux niveaux de pauvreté de la population, qu'une enquête fondée sur un échantillon de petite taille ne permet pas de détecter.

43. Pour pouvoir donner des informations sur la pauvreté à des niveaux plus localisés, il faut un ensemble de données provenant d'un échantillon de taille supérieure de plusieurs ordres de grandeurs à celui d'une enquête LSMS. Bien entendu, le plus gros ensemble de données de tout pays est le recensement de sa population. Tou-

³⁴ Voir Grosh (1997), par exemple.

³⁵ Pour plus de détails sur la méthodologie d'établissement de cartes de pauvreté fondées sur la mesure du bien-être telle qu'elle ressort des enquêtes, et de liens avec les données du recensement, voir : Hentschel *et al.* (2000); Elbers, Lanjouw et Lanjouw (2002; 2003); Elbers *et al.* (2001); et Demombynes *et al.* (2001). D'autres travaux sont en cours sur l'utilisation de cette technique pour lier les deux enquêtes; cependant, il est impossible d'estimer correctement les erreurs types résultant de l'établissement d'un tel lien.

³⁶ On peut utiliser d'autres enquêtes sur les ménages, à condition qu'elles donnent une solide mesure monétaire du bien-être, telle que la consommation totale ou le revenu total.

tefois, comme celui-ci couvre l'ensemble de la population, il ne recueille que des informations très limitées auprès de chaque ménage et n'a lieu généralement qu'une fois tous les 10 ans. Par conséquent, il ne peut servir à mesurer adéquatement la pauvreté. Un type nouveau de travail, qui permet de lier les données d'une enquête à celles du recensement, est à l'essai. Cette technique tire profit de la mesure du bien-être donnée par l'enquête LSMS et de la couverture assurée par le recensement. Elle implique l'estimation de la pauvreté à partir des données de l'enquête LSMS en utilisant un vecteur de variables que l'on trouve à la fois dans l'enquête et le recensement. On utilise ensuite les paramètres ainsi estimés avec les données du recensement pour prédire la probabilité de pauvreté de chaque ménage, et l'on établit des comptages de personnes pour de petites subdivisions du territoire national en utilisant les données du recensement. Les cartes de pauvreté ainsi établies fournissent un outil au gouvernement pour l'affectation des ressources. On trouve des exemples de telles cartes de pauvreté en Équateur, au Guatemala, à Madagascar, au Nicaragua, au Panama et en Afrique du Sud.

44. Le troisième exemple de l'utilisation des données des enquêtes LSMS pour améliorer le ciblage des programmes sociaux découle d'une évaluation du Fonds d'investissement social d'urgence au Nicaragua (connu sous son sigle espagnol de FISE). Cette évaluation traitait des questions de ciblage ainsi que de l'impact des investissements du FISE dans les domaines de l'approvisionnement en eau, des latrines, de l'éducation, de la santé et de l'assainissement³⁷. Pour ce faire, une enquête LSMS a été préparée à l'échelon national. Un suréchantillon de ménages était constitué de ménages des zones du projet du FISE ainsi que de communautés similaires situées hors des zones de ce projet. L'autre source de données était constituée par les dossiers du projet et les archives administratives qui ont été utilisés pour évaluer les coûts administratifs du projet.

45. Le suréchantillon de ménages des communautés de la zone du projet du FISE et de l'extérieur de cette zone a permis de créer à la fois des groupes de contrôle et des groupes d'analyse pour mesurer l'impact des investissements du FISE et l'efficacité de leur ciblage. Par ailleurs, l'échantillon national de l'enquête LSMS a permis de créer un deuxième groupe de contrôle (à l'aide des techniques de comparaison des propensions), ce qui a encore renforcé la validité de l'évaluation. L'évaluation de l'efficacité du ciblage s'est faite à la fois au niveau des communautés (les investissements du FISE étaient-ils progressifs dans les communautés où le projet était exécuté ?) et au niveau individuel (à l'intérieur des communautés bénéficiaires du projet du FISE, les segments les plus pauvres de la population avaient-ils plus de chances de bénéficier de l'investissement de ce fonds ?).

46. Cette évaluation a permis de montrer, avec des résultats statistiquement significatifs, l'efficacité d'ensemble du ciblage et d'évaluer les principaux types de projets sur la base des critères de ciblage. Elle a montré que les projets d'assainissement étaient fortement régressifs, tandis que les projets de latrines et d'éducation primaire étaient systématiquement progressifs, touchant les 17 % de la population classés comme étant en situation d'extrême pauvreté. Le résultat immédiat de l'évaluation a été la suspension des projets d'assainissement et la décision de faire porter les efforts sur l'amélioration des activités de diffusion et des investissements en faveur des communautés les plus pauvres. Le coût de cette évaluation très complexe du projet du FISE a représenté 1 % des investissements effectués par le projet jusqu'à la date de cette évaluation.

G. Conclusions

47. Les résultats des enquêtes LSMS ont montré l'intérêt de cette approche. Les données ont permis aux gouvernements de comprendre les effets de leurs politi-

³⁷ Pour plus de détails sur les buts de l'évaluation, les méthodes employées et les résultats, voir Banque mondiale (2000).

ques, de modifier celles-ci et de mieux cibler les ressources vers les groupes et les régions les plus pauvres. Cette importance attachée à la qualité a porté ses fruits, sous la forme de plus faibles niveaux d'erreurs et d'une utilité accrue des données. Cela dit, cette approche oblige à certains compromis. Ses coûts sont relativement élevés; la taille plus réduite de l'échantillon limite le niveau de fragmentation possible; et l'obligation d'une planification et d'une conception immédiates ralentit le processus; néanmoins, les données peuvent être produites rapidement, une fois le travail commencé, et les liens établis avec les décideurs permettent une utilisation accrue des données.

48. L'intégration des enquêtes LSMS au système d'enquêtes sur les ménages du pays présente clairement des avantages. La fréquence requise de ces enquêtes dépend toutefois de plusieurs facteurs. Premièrement, la décision d'effectuer une ou plusieurs enquêtes sur un certain laps de temps doit être fonction des besoins analytiques du pays. Si de nombreux programmes gouvernementaux peuvent être évalués à l'aide de données transversales (ciblage, incidence, voire impact à l'aide des techniques de comparaison de propension) des ensembles de données en coupes répétées et de données de panel sont nécessaires pour d'autres types d'analyse des changements dans le temps et de l'impact de certaines politiques et de certains événements.

49. Le deuxième facteur à prendre en considération, en ce qui concerne la fréquence d'exécution des enquêtes LSMS, a trait à la capacité d'analyse du pays. Les données ont besoin d'être analysées pour servir aux décideurs et pour que chaque nouvelle enquête marque un progrès par rapport aux précédentes. Si les données ne peuvent être analysées rapidement, une bonne part de l'investissement dans des cycles d'enquête répétées peut être perdue. En pareil cas, il peut être utile de laisser s'écouler un certain délai (par exemple, de trois ans) entre deux enquêtes.

50. Enfin, les questions budgétaires et logistiques sont souvent aussi importantes que les questions de fond dans le choix du moment ou de la périodicité des enquêtes. Autrement dit, la fréquence à laquelle une enquête se répète doit être fonction d'un compromis entre l'importance de ses résultats en regard de ceux d'autres enquêtes. Par ailleurs, il ne faut pas perdre de vue qu'aucune source de données ne peut répondre à tous les besoins. Les gouvernements ont besoin de dossiers administratifs et des données du système d'informatique de gestion des projets, ainsi que d'un ensemble d'enquêtes sur les ménages, tant pour leurs politiques macro que microéconomiques. Combinées à un ensemble complet d'enquêtes, les enquêtes LSMS peuvent favoriser une amélioration sensible de la compréhension de la façon dont la politique et les dépenses gouvernementales influent sur la vie de la population du pays.

Annexe I

Liste des Enquêtes pour l'étude sur la mesure des niveaux de vie

Pays	Année	Ménages interrogés
Albanie	1996	1 500
Albanie	2002	3 600
Arménie	1996	4 920
Azerbaïdjan	1995	2 016
Bolivie	1999	...
Bolivie	2000	5 032
Bolivie	2001	..
Bosnie-Herzégovine	2001	5 402
Brésil	1996-1997	4 940
Bulgarie	1995	2 500
Bulgarie	1997	2 317
Bulgarie	2001	2 633
Cambodge	1997	6 010
Chine : Hebei et Liaoning	1995 et 1997	780
Côte d'Ivoire	1985	1 588
Côte d'Ivoire	1986	1 600
Côte d'Ivoire	1987	1 600
Côte d'Ivoire	1988	1 600
Équateur	1994	4 500
Équateur	1995	5 500
Équateur	1998	5 801
Équateur	1998-1999	5 824
Gambie	1992	1 400
Ghana	1987-1988	3 200
Ghana	1988-1989	3 200
Ghana	1991-1992	4 565
Ghana	1998-1999	5 998
Guatemala	2000	7 276
Guinée	1994	4 705
Guyana	1992-1993	5 340
Inde : Uttar Pradesh et Bihar	1997-1998	2 250
Jamaïque	1988-2000 (annuelle)	2 000-7 300
Kazakhstan	1996	1 996
Kosovo	2000	2 880
Kirghizistan	1993	2 000
Kirghizistan	1996 (printemps)	..
Kirghizistan	1996 (automne)	1 951
Kirghizistan	1997	2 962
Kirghizistan	1998	2 979
Madagascar	1993	4 504
Malawi	1990	6 000

Pays	Année	Ménages interrogés
Mauritanie	1987	1 600
Mauritanie	1989	1 600
Mauritanie	1995	3 540
Maroc	1991	3 323
Maroc	1998	..
Népal	1996	3 373
Nicaragua	1993	4 200
Nicaragua	1998-1999	4 209
Nicaragua	2001	4 290
Niger	1989	1 872
Niger	1992	2 070
Niger	1995	4 383
Pakistan	1991	4 800
Panama	1997	4 945
Papouasie-Nouvelle-Guinée	1996	1 396
Paraguay	1997-1988	4 353
Paraguay	1999	5 101
Paraguay	2000-2001	8 131
Pérou	1985	5 120
Pérou (Lima seulement)	1990	1 500
Pérou	1991	2 200
Pérou	1994	3 500
Fédération de Russie ^a	1992	6 500
Afrique du Sud	1993	9 000
Tadjikistan	1999	2 000
République-Unie de Tanzanie : Kagera	1991-1994	840
République-Unie de Tanzanie : échelon national	1993	5 200
Tunisie	1995-1996	3 800
Ouganda	1992	9 929
Viet Nam	1992-1993	4 800
Viet Nam	1997-1998	5 999

Note : Deux points (..) indiquent l'absence de données.

^a L'Enquête longitudinale de suivi de 1992 en Russie a été financée par la Banque mondiale. La Banque mondiale n'a pas participé aux enquêtes suivantes. Pour plus de renseignements, voir le site Web du Carolina Population Center : http://www.cpc.unc.edu/projects/rlms/rlms_home.html.

Annexe II

Budget des enquêtes LSMS

Comme il est indiqué dans le texte du chapitre XXIII, il n'y a pas deux enquêtes LSMS absolument semblables, ni deux ONS, ni de similarité entre les coûts liés aux salaires, aux transports, aux équipements, etc. entre les différents pays. Il est donc impossible de donner des informations sur ce que peut coûter une telle enquête en un lieu et à un moment donné. Nous offrons donc ci-après un petit exercice destiné à aider le lecteur à commencer à établir un budget. Il s'agit simplement d'un guide rapide sur la façon d'estimer les coûts salariaux les plus élémentaires concernant le travail de terrain. En utilisant ce guide avec les coûts réels enregistrés dans le pays étudié, on peut obtenir une approximation très grossière de ce que pourrait coûter une enquête LSMS.

En moyenne, étant donné la complexité de l'instrument de l'enquête et l'utilisation de déclarants directs, l'enquêteur peut effectuer deux demi-entrevues par jour (voir figure XXIII.2 sur la façon dont l'enquête s'effectue). Si nous supposons une semaine de six jours de travail (que la journée libre soit prise chaque semaine ou étalée de quelque autre manière sur le mois), l'enquêteur peut effectuer 24 entrevues complètes par mois.

Supposons que nous ayons besoin d'un échantillon de 4 000 ménages. Si chaque enquêteur peut interroger 24 ménages par mois, il faut un total de 167 mois d'enquêteur pour interroger les 4 000 ménages. Si le travail de terrain est étalé sur une période de 12 mois, il faut donc 14 enquêteurs. Pour chaque ensemble de deux enquêteurs, il faut aussi un superviseur, un préposé à l'entrée des données, un chauffeur et un véhicule. Autrement dit, l'effectif total du personnel de terrain (sans compter la supervision régionale assurée par le personnel de l'ONS) comprend :

- 14 enquêteurs
- 7 superviseurs
- 7 préposés à l'entrée des données
- 7 chauffeurs

Si les planificateurs utilisent les paramètres indiqués ci-dessous, les coûts salariaux du travail de terrain nécessaire à l'enquête seront les suivants :

Membres de l'équipe	Coût par personne et par mois	Nombre de mois	Coût
14 enquêteurs	500	13	91 000
7 superviseurs	575	13,5	54 338
7 préposés à l'entrée des données	525	14	51 450
7 chauffeurs	300	13	27 300
Estimation grossière des coûts salariaux du travail de terrain			224 088

Note : Bien que le travail de terrain prenne 12 mois, on ajoute un mois supplémentaire pour couvrir les coûts de la formation (pendant laquelle le personnel de terrain reçoit généralement une certaine rémunération) et/ou les retards dans le travail d'enquête. Les préposés à l'entrée des données sont souvent gardés un mois de plus pour parachever le travail et, s'il y a lieu, nettoyer l'ensemble des données.

D'après la figure XXIII.3, les coûts du personnel de terrain, qui représentent les trois quarts du total des coûts salariaux de l'enquête, ne représentent eux-mêmes que 28 % des coûts totaux de l'enquête. Par conséquent, en se fondant sur un simple calcul, on parvient à une estimation approximative de 1 067 086 du coût total de l'enquête.

Il est clair que ce chiffre n'est qu'une très grossière approximation. On a également besoin de certains détails de coût tels que ceux de l'assistance technique, par exemple. Toutefois, ce simple exercice initial peut être utile pour démarrer le processus d'établissement du budget de l'enquête. Le lecteur est invité à se reporter au chapitre 8 de Grosh et Muñoz (1996) pour un exposé détaillé de la façon de concevoir un budget réaliste pour une enquête LSMS.

Annexe III

Effet de la conception de l'échantillon sur la précision et l'efficacité des enquêtes LSMS¹

¹ La présente annexe s'inspire pour beaucoup de précédents travaux de Temesgen et Morganstein (2000).

A. Introduction

D'autres chapitres de cette publication fournissent des informations détaillées sur les questions d'échantillonnage et, en particulier, sur l'effet d'échantillons de conception complexe ou à phases multiples sur la variance des estimations obtenues. Cet effet « de conception » est commun à toutes les enquêtes qui n'utilisent pas un échantillon aléatoire simple, telles que les enquêtes LSMS. Il constitue un élément de l'erreur globale d'échantillonnage : la différence entre une estimation obtenue à partir d'un échantillon en grappes à phases multiples et une estimation qui serait obtenue à l'aide d'un échantillon aléatoire de conception simple. Dans la présente annexe, nous résumons les principales questions et nous nous efforçons de montrer l'impact effectif de la conception de l'échantillon sur plusieurs enquêtes LSMS.

B. Calcul des erreurs d'échantillonnage, des effets de conception et de certains éléments connexes

Dans un simple échantillon aléatoire, toutes les unités sélectionnées ont une probabilité de sélection identique et indépendante. Toutefois, on n'utilise presque jamais l'échantillonnage aléatoire simple pour les enquêtes sur les ménages, en raison de préoccupations logistiques et de considérations de coût. À la place, comme dans les enquêtes LSMS, on utilise des échantillons de conception plus complexe, à phases multiples, comportant une stratification ou une mise en grappes. Cela influe sur le calcul de la variance des estimations et sur l'efficacité de l'échantillon lui-même. Pour calculer les erreurs d'échantillonnage pour les échantillons à phases multiples, il faut connaître les variables qui définissent les strates, les unités primaires d'échantillonnage (UPE) et les procédures de pondération (éventuelles) utilisées pour leur conception. Lorsque ces variables sont identifiées, on peut utiliser un certain nombre de logiciels statistiques pour calculer les mesures nécessaires².

Les mesures des erreurs d'échantillonnage indiquées ici pour certaines variables au niveau des ménages et au niveau des individus dans les enquêtes LSMS comprennent l'erreur type [*standard error (SE)*] que l'on calcule en tenant compte de la complexité de conception de l'échantillon, du coefficient de variation [*CV (%)*], de la taille de l'échantillon (*n*), de l'effet de conception, du coefficient de corrélation intra-classe (ρ), des limites inférieure et supérieure des intervalles de confiance (*CI*), et de la taille effective de l'échantillon (*EFFn*). Ces termes sont tous définis aux chapitres II, VI, VII et dans d'autres chapitres.

C. Erreurs types, effets de conception et corrélation intra-classe calculés à partir des enquêtes LSMS

L'un des aspects importants du calcul des erreurs d'échantillonnage pour les variables d'une enquête est qu'il permet de comparer l'efficacité (la précision) de différents types de conception de l'échantillon entre eux et avec la précision qui aurait été obtenue avec un échantillon aléatoire simple de la même taille. Outre le fait qu'il indique la fiabilité des données de l'enquête, cet exercice permet aux analystes d'évaluer

² On a utilisé le logiciel de statistiques Wesvar pour ces calculs. Parmi les autres programmes qui peuvent servir à estimer les variances de l'échantillonnage et tout un ensemble de statistiques connexes utilisées pour les enquêtes de conception complexe, on peut citer : CENVAR, CLUSTERS, Epi-Info, PC-CARP, SUDAAN, VPLX et STATA. Certains de ces logiciels peuvent être téléchargés gratuitement à partir de l'Internet.

l'efficacité d'une conception particulière et de guider la conception des enquêtes à venir. Les trois tableaux ci-dessous comparent les effets de conception et certaines mesures connexes pour plusieurs variables afin de montrer les différences qui existent : a) à l'intérieur d'un même pays entre différentes variables; b) à l'intérieur d'un même pays à différents moments; et c) entre pays³.

Comme le montre le tableau AIII.1, à l'intérieur d'un même pays, la même enquête engendre des effets de conception sensiblement différents pour différentes variables. Ce tableau repose sur les données de l'enquête LSMS effectuée en 1987 au Ghana et sur les variables établies aux niveaux des ménages et des individus. Pour certaines variables, telles que les dépenses totales par personne, où la corrélation intra-classe est faible, l'effet de conception n'est pas élevé (1,9); en revanche, pour d'autres variables, telles que l'accès à l'assainissement et à l'eau, où les corrélations intra-classe sont élevées (l'infrastructure tend à être concentrée sur certains endroits particuliers), les effets de conception sont élevés (7,8 et 8,0 respectivement); ces effets sont encore plus élevés pour certaines sous-catégories de population urbaine ou rurale.

³ Pour le rapport complet, se reporter à Temesgen et Morganstein (2000).

Tableau AIII.1
Variation des effets de conception par variable, Ghana, 1987

Variable		Estimation	SE	CV (%)	Intervalle de confiance		n	Effet de conception	EFFn	ρ
					Limite inférieure	Limite supérieure				
Accès à l'électricité	Total	0,267	0,019	7,265	0,229	0,305	3 138	6,034	520	0,300
	Rural	0,078	0,022	28,744	0,034	0,121	2 023	14,063	144	0,787
	Urbain	0,611	0,041	6,714	0,530	0,691	1 115	7,888	141	0,403
Taille du ménage	Total	4,940	0,083	1,682	4,777	5,103	3 136	2,089	1 501	0,065
	Rural	5,147	0,097	1,877	4,958	5,336	2 022	1,735	1 165	0,044
	Urbain	4,565	0,165	3,615	4,241	4,888	1 114	3,291	339	0,134
Propriété foncière	Total	0,591	0,024	4,018	0,544	0,637	3 138	7,315	429	0,376
	Rural	0,747	0,033	4,393	0,683	0,811	2 023	11,520	176	0,634
	Urbain	0,308	0,035	11,413	0,239	0,376	1 115	6,453	173	0,319
Dépenses totales par personne	Total	82 745,2	1 902,2	2,3	79 017,1	86 473,4	3 104	1,883	1 648	0,053
	Rural	70 908,1	2 526,4	3,6	65 956,3	75 859,8	2 001	3,100	646	0,127
	Urbain	104 219,5	3 702,1	3,6	96 963,6	111 475,4	1 103	1,759	627	0,044
Dépenses alimentaires par personne	Total	56 779,3	1 309,2	2,3	54 213,2	59 345,3	3 104	1,927	1 611	0,055
	Rural	52 382,3	1 777,9	3,4	48 897,6	55 867,0	2 001	2,577	776	0,095
	Urbain	64 756,0	2 147,9	3,3	60 546,2	68 965,8	1 103	1,580	698	0,034
Enlèvement salubre des ordures	Total	0,019	0,003	16,647	0,013	0,026	3 135	1,724	1 818	0,043
	Rural	0,010	0,003	29,044	0,004	0,016	2 020	1,704	1 185	0,042
	Urbain	0,037	0,009	23,481	0,020	0,054	1 115	2,347	475	0,079
Accès à des toilettes salubres	Total	0,590	0,025	4,159	0,542	0,638	3 135	7,808	401	0,405
	Rural	0,659	0,034	5,091	0,593	0,725	2 020	10,114	200	0,549
	Urbain	0,465	0,038	8,092	0,392	0,539	1 115	6,357	175	0,313
Accès à l'eau potable	Total	0,395	0,025	6,251	0,347	0,443	3 135	7,994	392	0,416
	Rural	0,224	0,031	13,818	0,164	0,285	2 020	11,150	181	0,611
	Urbain	0,704	0,046	6,482	0,615	0,793	1 115	11,144	100	0,593

Source : Temesgen et Morganstein (2000).

Note : Pour des descriptions des variables utilisées, voir tableaux AIII.4 et AIII.5 ci-après.

Le tableau AIII.2, également établi à partir de données recueillies au Ghana, montre que les effets de conception peuvent varier dans le temps au même titre qu'en fonction des variables. Ici, l'écart entre les deux enquêtes n'est que d'un an et la conception de base de l'échantillon n'a pas changé, mais les effets de conception, eux, ont varié : l'estimation concernant l'accès à la santé est devenue nettement plus précise (l'effet de conception est tombé de 5,01 à 3,64) et l'effet de conception concernant le chômage a aussi diminué, quoique dans une moindre mesure. L'autre variable du tableau, l'alphabétisme des adultes, a été mesurée avec moins de précision lors de la deuxième année de l'enquête.

Tableau AIII.2
Variation des effets de conception dans le temps, Ghana, 1987 et 1988

Ghana, 1987										
Variable		Estimation	SE	CV (%)	Intervalle de confiance		n	Effet de conception	EFFn	ρ
					Limite inférieure	Limite supérieure				
Alphabétisme des adultes	Femmes	0,402	0,021	5,103	0,362	0,442	1 339	2,342	572	0,080
	Hommes	0,613	0,018	2,953	0,578	0,649	1 381	1,910	723	0,054
	Total	0,509	0,016	3,192	0,477	0,541	2 720	2,875	946	0,112
Accès aux services de santé	Femmes	0,443	0,016	3,625	0,411	0,474	2 756	2,876	958	0,112
	Hommes	0,423	0,017	4,017	0,390	0,457	2 542	3,011	844	0,120
	Total	0,433	0,015	3,517	0,403	0,463	5 298	5,013	1 057	0,239
Chômage	Femmes	0,039	0,004	10,063	0,031	0,047	4 011	1,655	2 424	0,039
	Hommes	0,047	0,004	9,136	0,038	0,055	3 543	1,454	2 437	0,027
	Total	0,043	0,003	7,666	0,036	0,049	7 554	1,983	3 810	0,059

Ghana, 1988										
Variable		Estimation	SE	CV (%)	Intervalle de confiance		n	Effet de conception	EFFn	ρ
					Limite inférieure	Limite supérieure				
Alphabétisme des adultes	Femmes	0,390	0,022	5,526	0,348	0,432	1 289	2,519	512	0,090
	Hommes	0,587	0,020	3,397	0,548	0,626	1 226	2,013	609	0,060
	Total	0,486	0,018	3,654	0,451	0,521	2 515	3,179	791	0,130
Accès aux services de santé	Femmes	0,375	0,013	3,558	0,348	0,401	2 921	2,215	1 319	0,072
	Hommes	0,365	0,015	4,118	0,335	0,394	2 606	2,539	1 026	0,092
	Total	0,370	0,012	3,346	0,346	0,394	5 527	3,635	1 521	0,157
Chômage	Femmes	0,036	0,003	9,593	0,029	0,042	3 852	1,307	2 946	0,018
	Hommes	0,034	0,003	9,885	0,027	0,041	3 260	1,123	2 904	0,007
	Total	0,035	0,003	7,306	0,030	0,040	7 112	1,372	5 185	0,022

Source : Temesgen et Morganstein (2000),

Note : Pour des descriptions des variables utilisées, voir tableaux AIII.4 et AIII.5 ci-après.

Enfin, comme prévu, les effets de conception varient considérablement d'un pays à un autre. Le tableau AIII.3 montre comme les enquêtes effectuées en Côte d'Ivoire et au Pakistan ont produit des effets de conception très différents pour les mêmes variables. Ce résultat est dû à la fois à des conceptions différentes de l'échantillon et aux différences de caractéristiques entre les deux pays.

Tableau AIII.3
Variation des effets de conception entre pays

Côte d'Ivoire, 1988										
Variable		Estimation	SE	CV (%)	Intervalle de confiance		n	Effet de conception	EFFn	ρ
					Limite inférieure	Limite supérieure				
Alphabétisme des adultes	Total	0,567	0,031	5,538	0,506	0,629	1 660	6,676	249	0,378
	Rural	0,411	0,042	10,212	0,329	0,493	745	5,415	138	0,294
	Urbain	0,738	0,024	3,217	0,691	0,784	915	2,661	344	0,111
Accès aux services de santé	Total	0,417	0,029	6,883	0,361	0,473	1 849	6,260	295	0,351
	Rural	0,303	0,034	11,174	0,236	0,369	1 051	5,693	185	0,313
	Urbain	0,622	0,025	4,078	0,572	0,671	798	2,181	366	0,079
Chômage	Total	0,038	0,007	18,837	0,024	0,052	4 979	6,991	712	0,399
	Rural	0,007	0,003	50,457	0,000	0,013	2 529	4,357	580	0,224
	Urbain	0,081	0,013	16,218	0,055	0,107	2 450	5,679	431	0,312

Pakistan, 1991										
Variable		Estimation	SE	CV (%)	Intervalle de confiance		n	Effet de conception	EFFn	ρ
					Limite inférieure	Limite supérieure				
Alphabétisme des adultes	Total	0,5	0,013	2,5	0,48	0,53	6 834	4,335	1 577	0,222
	Rural	0,42	0,017	3,95	0,39	0,45	3 249	3,669	885	0,178
	Urbain	0,68	0,018	2,616	0,64	0,71	3 585	5,156	695	0,277
Accès aux services de santé	Total	0,5	0,012	2,329	0,48	0,52	9 238	5,02	1 840	0,268
	Rural	0,46	0,015	3,177	0,43	0,49	4 752	4,048	1 174	0,203
	Urbain	0,61	0,017	2,74	0,57	0,64	4 486	5,185	865	0,279
Chômage	Total	0,03	0,003	9,735	0,02	0,03	18 232	4,633	3 935	0,242
	Rural	0,02	0,003	14,955	0,02	0,03	8 934	4,706	1 898	0,247
	Urbain	0,03	0,003	8,956	0,03	0,04	9 298	2,539	3 662	0,103

Source : Temesgen et Morganstein (2000),

Note : Pour des descriptions des variables utilisées, voir tableaux AIII.4 et AIII.5 ci-dessous.

En résumé, la petite taille des échantillons utilisés dans les enquêtes LSMS et le fait que ces échantillons sont à phases multiples comporte des inconvénients du point de vue de la précision des estimations de l'échantillon. Par exemple, l'effet de conception pour l'alphabétisme des adultes hommes et femmes était élevé, à 6,7, dans l'enquête de 1988 en Côte d'Ivoire. Cela signifie que la précision de l'estimation avec un échantillon de taille (n) de 1 660 équivaut à celle obtenue avec un échantillon SRS de 249 seulement. En revanche, si nous ne considérons que le milieu urbain, nous constatons que l'effet de conception est légèrement plus faible (2,7), quoique toujours supérieur à 1, ce qui signifie qu'un échantillon de 915 personnes a la même précision qu'un échantillon SRS de 344 personnes. Devant l'ampleur possible de l'effet de conception et les écarts de taille de ces effets selon les variables, le temps et les pays, il est essentiel que les analystes tiennent compte de la conception de l'échantillon lorsqu'ils utilisent les données et, en particulier, lorsqu'ils effectuent des tests de signification statistique. Cela montre également qu'il est difficile de concevoir des échantillons efficaces pour les enquêtes à thèmes multiples sur les ménages. En essayant de réduire l'effet de conception d'une variable, on risque fort d'accentuer cet effet pour d'autres variables.

La solution empirique à ce problème est donc, dans la mesure du possible, de considérer essentiellement la (les) variable(s) importante(s) pour l'enquête.

Tableau AIII.4

Description des variables de l'analyse : niveau individuel

Variable	Description	Population étudiée
Chômage	Adultes actuellement sans emploi mais disponibles et à la recherche d'un emploi.	Personnes âgées de 15 à 64 ans
Accès aux services de santé	Proportion des personnes qui ont été malades durant le mois ayant précédé l'entrevue et qui se sont rendues dans des établissements de santé modernes tels qu'hôpitaux, cliniques et dispensaires (mais n'ont pas consulté de sages-femmes, de guérisseurs ou autres dispensateurs de soins coutumiers).	Personnes ayant été malades durant le mois précédent
Alphabétisme des adultes	Proportion d'adultes sachant lire (définis comme ceux qui peuvent lire un journal).	Personnes âgées de 15 à 24 ans

Tableau AIII.5

Description des variables de l'analyse : niveau des ménages

Variable	Description
Accès à une eau salubre	Proportion de ménages ayant accès à l'eau potable. Au niveau des ménages, cette variable est égale à l'unité si le ménage tire son eau potable d'un robinet, d'une conduite ou d'un puits équipé d'une pompe. Elle est nulle si la source d'approvisionnement en eau — cours d'eau, canal, lac ou marécage — est jugée potentiellement dangereuse pour la santé.
Propriété foncière	Proportion de ménages possédant une terre. Pour chaque ménage, cette variable est égale à l'unité si le ménage possède une terre, et à zéro dans le cas contraire.
Accès à l'électricité	Proportion de ménages ayant accès à l'électricité. Pour chaque ménage, cette variable est égale à l'unité si le ménage a accès à l'électricité, et à zéro dans le cas contraire.

RÉFÉRENCES

- Banque mondiale (2000), *Nicaragua: Ex-Post Impact Evaluation of the Emergency Social Investment Fund (FISE)*, Report No. 20400-NI. Washington, D.C.
- _____ (2001), *Jamaica Survey of Living Conditions (JLSC) 1988-98: Basic Information*. Washington, D.C., Groupe de recherche sur l'économie du développement.
- _____ (2002a), *Basic Information Document: Bosnia and Herzegovina Living Standards Measurement Study Survey*. Washington, D.C., Groupe de recherche sur l'économie du développement, Banque mondiale.
- _____ (2002b), *Guatemala Poverty Assessment*, Report No. 24221-GU. Washington, D.C.
- _____ (2002c), *The 1993 Nicaragua Living Standards Measurement Survey: Documentation*. Washington, D.C., Groupe de recherche sur l'économie du développement, Banque mondiale.
- Blank, Lorraine, et Margaret Grosh (1999). Using household surveys to build analytic capacity, *The World Bank Research Observer*, vol. 14, n° 2 (août), pp. 209-227.
- Demombynes, Gabriel, *et al.* (2001). *Producing an Improved Geographic Profile of Poverty: Methodology and Evidence from Three Developing Countries*, WIDER Discussion Paper, No. 2002/39. Helsinki, World Institute for Development Economic Research/ Université des Nations Unies.
- Elbers, C. J., Lanjouw et P. Lanjouw (2002). *Micro-level Estimation of Welfare*, Policy Research Working Paper, No. 2911. Washington, D.C., Banque mondiale.
- _____ (2003). Micro-level estimation of poverty and inequality. Dans *Econometrica*, vol. 71, n° 1, pp. 355-364,

- _____, *et al.* (2001). *Poverty and Inequality in Brazil: New Estimates from Combined PPV-PNAD Data*. Washington, D.C., Groupe de recherche sur l'économie du développement, Banque mondiale.
- Feres, Juan Carlos (1998). Falta de respuesta a las preguntas sobre el ingreso: su magnitud y efectos en las Encuestas de Hogares en América Latina. Dans *Conference Proceedings from the 2^o Taller Regional del Medición del Ingreso en las Encuestas de Hogares*, Buenos Aires, November, Santiago du Chili, Commission économique pour l'Amérique latine et les Caraïbes (CEPALC). Document LC/R.1886.
- Ghana Statistical Service (2000). *Ghana Living Standards Survey, Round Four (GLSS 4) 1998-99: Data User's Guide*. Accra.
- Grosh, Margaret (1991). *The Household Survey as a Tool for Policy Change: Lessons from the Jamaican Survey of Living Conditions*, Living Standards Measurement Study Working Papers, No. 80. Washington, D.C., Banque mondiale.
- _____. (1997). The policymaking uses of multi-topic household survey data: a primer, *The World Bank Research Observer*, vol, 12, n° 2, pp. 137-60,
- _____, et Paul Glewwe (1995). *A Guide to Living Standards Measurement Study Surveys and Their Data Sets*, Living Standards Measurement Study Working Paper, No. 120. Washington, D.C., Banque mondiale.
- _____, eds, (2000). *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 Years of the Living Standards Measurement Study Surveys*. Washington, D.C., Banque mondiale.
- _____, et Juan Muñoz (1996). *A Manual for Planning and Implementing the Living Standards Measurement Study Survey*. Living Standards Measurement Study Working Paper, No. 126. Washington, D.C., Banque mondiale.
- Hentschel, J., *et al.* (2000). Combining household data with census data to construct a disaggregated poverty map: a case study of Ecuador. *World Bank Economic Review*, vol, 14, n° 1 (janvier).
- Kish, Leslie (1965). *Survey Sampling*. New York, John Wiley and Sons, Inc.
- Pradhan, Menno, et Martin Ravallion (2000). Measuring poverty using qualitative perceptions of consumption adequacy, *Review of Economics and Statistics*, vol. 82, pp. 462-471.
- Ravallion, Martin, et Michael Lokshin (2001). Identifying welfare effects using subjective questions. *Economica*, vol. 68, pp. 335-357.
- _____. (2002). Self-rated economic welfare in Russia. *European Economic Review*, sous presse.
- Ryten, Jacob (2000), The MECOVI Program: ideas for the future: a mid-term evaluation. Document non publié, préparé pour la Banque interaméricaine de développement. Décembre.
- Skinner, C. J. D. Holt et T. M. F. Smith (1989), *Analysis of Complex Surveys*. Chichester, Royaume-Uni, John Wiley and Sons.
- Temesgen, Tilahun, et David Morganstein (2000), *Measurement of Sampling Errors: Application to Selected Variables in LSMS Surveys*. Washington, D.C., Groupe de recherche sur l'économie du développement, Banque mondiale.

Chapitre XXIV

Conception de l'enquête et de l'échantillon pour les enquêtes sur le budget des ménages

HANS PETTERSSON
 Statistics Sweden
 Stockholm, Suède

RÉSUMÉ

Le présent chapitre traite de certaines questions concernant la conception de l'enquête et de l'échantillon pour les enquêtes sur le budget des ménages, et plus particulièrement dans les pays en développement. La section B examine de façon détaillée les problèmes de mesure de la consommation et des revenus. La section C traite de certaines questions cruciales concernant la conception de l'échantillon, notamment de la stratification et de l'étalement de l'échantillon dans l'espace (répartition géographique) et dans le temps (sur l'ensemble de la saison). La section D donne une description de l'Enquête de 1997/98 sur les dépenses et la consommation en République démocratique populaire lao (LECS-2), et la section E analyse certaines des leçons tirées de l'expérience de cette enquête.

Termes clés : enquête sur le budget des ménages, enquête sur les dépenses et la consommation, mesure des dépenses, méthode du journal.

A. Introduction

1. L'appellation « Enquête sur le budget des ménages » est purement générique et recouvre une large catégorie d'enquêtes. Il peut s'agir « d'enquêtes sur les dépenses des familles », « d'enquêtes sur les dépenses et la consommation » ou « d'enquêtes sur le revenu et les dépenses », mais le trait commun est un effort pour saisir les éléments importants du « budget » quotidien des ménages. Certaines enquêtes, conçues à l'origine comme « enquêtes sur le budget des ménages » sont devenues des enquêtes à buts multiples. À l'essentiel des questions sur la consommation, les dépenses et le revenu des ménages sont venus d'ajouter des modules supplémentaires traitant, par exemple, de la santé, de la nutrition et de l'éducation. Cette façon d'intégrer plusieurs thèmes à une enquête à buts multiples se répand de plus en plus. Dans le présent chapitre, nous nous concentrerons sur les enquêtes dont l'un des éléments importants est la mesure du budget des ménages, qu'il s'agisse d'une enquête à buts multiples ou d'une étude plus spécifiquement axée sur le budget.

2. Les données sur la consommation, les dépenses et le revenu des ménages peuvent servir à divers usages. Elles peuvent être utilisées pour diverses études des caractéristiques socio-économiques de la population et de la façon dont elles sont réparties (par exemple, étude de la prévalence de la pauvreté). Lorsque ces enquêtes sont effectuées régulièrement, elles permettent de suivre la condition de divers groupes de population. Les enquêtes effectuées par la Banque mondiale pour l'Étude sur la mesure des niveaux de vie [*World Bank Living Standard Measurement Study (LSMS)*]

étaient conçues spécialement pour mesurer la pauvreté et les différences de niveau de vie au sein de la population. Ces dernières années, ces enquêtes ont beaucoup servi à mesurer les résultats des interventions gouvernementales, notamment les effets des projets de lutte contre la pauvreté. Ces données peuvent également servir à façonner les décisions des gouvernements dans les domaines sociaux et budgétaires.

3. Les données des enquêtes sur le budget des ménages représentent un apport important pour le système national de statistiques économiques, et en particulier pour la comptabilité nationale. Ces enquêtes permettent de mesurer la consommation des ménages, ainsi que la production des entreprises et des activités agricoles familiales (qui représentent une part importante de la production nationale des pays pauvres). Dans le système de statistiques économiques, on met l'accent sur les agrégats nationaux. Par conséquent, une enquête axée principalement sur les besoins de ce système devrait être conçue de manière à donner des estimations des totaux à l'échelon national. Dans certains cas, une telle conception peut être moins efficace lorsque l'on utilise les données de l'enquête pour une analyse visant la conception de politiques gouvernementales ou l'évaluation des interventions, où l'on cherche à déterminer les différences entre divers groupes de population ou diverses régions géographiques.

4. Dans ce chapitre, nous examinerons les enquêtes sur le budget des ménages dans les pays en développement en tant que sources de données pour le système de statistiques économiques. Ce chapitre compte quatre grandes sections. La section B traite de certains problèmes importants concernant la conception de l'enquête, en particulier des difficiles problèmes de mesure, notamment de mesure de la consommation des ménages. La section C analyse des questions de conception de l'échantillon dans les enquêtes sur le budget des ménages. La section D est une étude de cas sur l'Enquête sur les dépenses et la consommation en République démocratique populaire lao, 1997/98 (LECS-2). Et la section E contient un examen des expériences et des leçons à tirer de cette enquête.

B. Conception de l'enquête

1. Méthodes de collecte de données utilisées dans les enquêtes sur le budget des ménages

5. Les enquêtes sur le budget des ménages ont pour principal but de mesurer la consommation totale des ménages et ses composants. L'approche traditionnelle du problème de la mesure, qui est encore suivie dans de nombreuses enquêtes, est celle d'une collecte d'informations détaillées. Il est demandé aux ménages de rendre compte séparément d'un grand nombre de leurs achats, tant en quantités achetées et en unités monétaires. Une autre approche tend à limiter la collecte de données de consommation à un niveau moins détaillé. Telle est l'approche généralement suivie dans les Enquêtes de la Banque mondiale sur la mesure des niveaux de vie (Deaton, 1997).

6. Il existe essentiellement deux moyens de recueillir les données sur la consommation :

- Par des entrevues avec les ménages au cours desquelles leur sont posées des questions sur leur consommation;
- Par la tenue d'un journal dans lequel les ménages enregistrent quotidiennement leurs dépenses de consommation et autres.

7. La méthode du journal nécessite au moins deux visites au ménage : une au début et une à la fin de la période de tenue du journal. Souvent, on prévoit une visite à mi-parcours pour s'assurer que le journal est bien tenu. L'entrevue rétrospective peut

faire l'objet d'une seule visite au ménage, mais il est courant qu'elle donne lieu à deux visites.

2. Problèmes de mesure

8. Comment la consommation d'un ménage doit-elle se mesurer lors d'une entrevue où sont posées des questions à caractère rétrospectif ? Faut-il la mesurer de façon détaillée pour un grand nombre d'articles ou de façon moins détaillée ? La première approche donne des informations plus détaillées que la seconde, mais moyennant un coût sensiblement plus élevé. Si nous pouvons nous passer d'un tel niveau de détail et ne cherchons qu'à estimer la consommation totale, la seconde approche permet-elle de produire des estimations aussi précises que la première ? Il n'y a pas de conclusion définitive concernant la précision. Deaton mentionne des études effectuées ces dernières années, notamment une enquête pilote en Indonésie portant sur 8 000 ménages, qui a permis de tester deux questionnaires (Deaton, 1997). Un questionnaire long avait des questions sur 218 produits alimentaires et 102 produits non alimentaires, et le questionnaire court en avait sur 15 produits alimentaires et 8 produits non alimentaires. Les estimations du total des dépenses alimentaires découlant des réponses à ces deux questionnaires différaient légèrement. Les estimations des dépenses non alimentaires étaient d'environ 15 % plus élevées avec le questionnaire long (Banque mondiale, 1992, appendice 4.2). Toutefois, d'autres tests n'ont pas donné les mêmes résultats. Des tests similaires effectués en El Salvador (Joliffe et Scott, 1995) et à la Jamaïque (*Statistical Institute and Planning Institute of Jamaica*, 1996, appendice III) laissent apparaître des différences plus sensibles entre les deux questionnaires. Dans le cas d'El Salvador, le long questionnaire affichait des dépenses totales de 40 % plus élevées et des dépenses non alimentaires de 27 % plus élevées. Quant au test effectué à la Jamaïque, il affichait des chiffres de 26 % plus élevés sur le questionnaire long pour les dépenses totales. La conclusion de Deaton a été que : « Bien que le questionnaire plus court permette parfois une grosse économie sur les coûts et la durée — en Indonésie, de 80 à 10 minutes — de l'enquête, il semble que cette économie se fasse aux dépens de la précision » (Deaton, 1997).

9. La méthode du journal réduit les risques que comporte le fait de se fier à la mémoire des répondants. Cela dit, cette méthode est difficile à appliquer quand une large fraction de la population est illettrée. Même avec un taux élevé d'alphabétisme, il faut s'attendre à des difficultés avec cette méthode, en ce sens que les ménages plus pauvres sont moins à même d'utiliser un journal et que nombre des ménages qui le peuvent omettent souvent d'y recourir (Deaton et Grosh, 2000). Lors de l'Enquête de 1995 à buts multiples sur les ménages, le Bureau général de la statistique du Viet Nam a constaté que, dans les zones urbaines, beaucoup de ménages ne tenaient pas leur journal (Glewwe et Yansaneh, 2001). La durée pendant laquelle ce journal doit être tenu est également une question à considérer, cette période étant de deux semaines pour de nombreuses enquêtes et d'un mois complet pour d'autres. La recherche a montré une tendance à indiquer moins de dépenses pendant la deuxième semaine que pendant la première pour les journaux à tenir pendant deux semaines, ce qui dénote probablement un effet de fatigue.

10. Beaucoup d'enquêtes sur le budget des ménages collectent également des données sur leur revenu. La mesure du revenu des ménages pose un problème encore plus difficile que celle de leur consommation. Le revenu est un sujet délicat pour beaucoup de répondants, surtout parmi les populations aisées. Les répondants ont parfois tendance à craindre que les informations sur leurs revenus ne servent à des fins fiscales, surtout lorsque existe une entreprise familiale.

11. Les revenus doivent être enregistrés pour tous les membres de la famille et pour tous les types de revenus (revenus tirés d'une entreprise familiale ou de l'agriculture, revenus informels provenant d'activités à temps partiel, revenus des actifs, etc.). Les calculs des revenus sont encore compliqués par les dons en espèces et en nature, les envois de fonds des travailleurs émigrés et les prêts. Les revenus agricoles des familles de petits exploitants posent des problèmes particuliers, car ces familles tirent une partie de leur alimentation de leur production de subsistance. Par ailleurs, une partie de leur revenu monétaire peut provenir de la vente occasionnelle de produits agricoles et se révéler difficile à chiffrer avec précision lors de l'entrevue.

12. Il est probable et, dans certains cas, prouvé, que du fait de ces difficultés théoriques et pratiques d'évaluation du revenu des ménages, celui-ci tend à être sous-estimé. L'expérience des enquêtes sur les revenus et les dépenses tendant à corroborer ce fait. On constate souvent que les estimations du revenu sont sensiblement plus faibles que celles de la consommation, à tel point qu'il est difficile d'expliquer la différence par l'utilisation par les ménages de leur épargne pour couvrir leur consommation. L'autre explication possible – à savoir que la consommation est surestimée – est plus contestable. La recherche indique que la consommation risque plus d'être sous-estimée que surestimée. On a donc des raisons de penser que nombre d'estimations du revenu sont trop faibles.

3. Périodes de référence

13. La décision concernant la période de référence est étroitement liée à la décision relative à l'instrument de mesure (le questionnaire « long » ou « court », la méthode du journal pour la consommation alimentaire ou les questions de mémoire, etc.). La période de référence qu'il est demandé au répondant de se rappeler ne doit pas être trop longue, car cela augmenterait les risques d'erreurs. L'effet du prolongement de la période de référence a été étudié dans une expérience effectuée dans le cadre de l'Enquête sur les niveaux de vie au Ghana. Cette étude a montré que, pour 13 articles fréquemment achetés, les dépenses signalées diminuaient en moyenne de 2,9 % par journée ajoutée à la période de référence (Scott et Amenuvegebe, 1990). Les avis sont partagés parmi les chercheurs sur les effets de diverses périodes de référence. Une étude précédente effectuée sur l'Enquête nationale par sondage en Inde semble indiquer que, pour certains aliments, une période de référence d'un mois engendre moins de distorsion qu'une période d'une semaine (Mahalanobis et Sen, 1954). Les Enquêtes de ces dernières années sur la mesure du niveau de vie semblent confirmer les résultats de Scott, mais on ne voit pas clairement si ces résultats sont dus aux défaillances de mémoire sur les périodes plus longues ou au télescopage des données sur les périodes plus courtes (Deaton, 1997).

14. Les articles achetés fréquemment, comme les aliments, ont généralement des périodes de référence plus courtes, tout au plus d'un mois. Pour les articles dont les achats sont moins fréquents, la situation est différente. Le rappel de dépenses sur des articles aux achats moins fréquents, tels que les biens durables, doit porter sur une période relativement plus longue, parce qu'une période trop courte entraînerait de fortes variances dans les estimations des totaux. La longueur de la période de référence adéquate diffère donc sensiblement selon des types d'articles.

4. Fréquence des visites

15. Dans la plupart des enquêtes sur les revenus et les dépenses, les données sont recueillies à l'occasion de visites répétées aux ménages de l'échantillon. La fréquence de ces visites à chaque ménage dépend de la méthode de mesure. La procédure

normale suivie pour la méthode rétrospective prévoit deux visites, espacées plus ou moins de deux semaines. Dans les enquêtes utilisant la méthode du journal, on recommande une ou deux semaines d'intervalle entre les visites de suivi aux ménages.

16. Des visites répétées à la même famille peuvent provoquer une certaine fatigue chez les répondants et une détérioration de la qualité de leurs réponses. Il faut donc trouver le juste équilibre entre les avantages que comporte le fait de suivre plus longtemps le ménage et la fatigue que cela peut lui imposer.

17. L'autre type d'enquête à visites répétées est celui où le ménage est interrogé pendant deux périodes de référence ou plus étalées sur l'année. Un exemple de ce type d'enquête nous est offert par l'Enquête de 1995/96 sur le revenu, la consommation et les dépenses des ménages en Éthiopie, où les ménages ont reçu deux visites durant deux saisons différentes et ont été interrogés sur le mois précédant chaque visite. Cette situation est examinée de façon plus approfondie dans la section sur l'échantillonnage ci-dessous.

5. Non-réponse

18. L'un des traits distinctifs des enquêtes sur le budget des ménages est le poids imposé aux ménages de l'échantillon par les réponses aux questions. Le taux de refus est généralement plus élevé dans ces enquêtes que dans les autres, et il peut être très élevé parmi certains segments de la population. Aux refus initiaux viennent s'ajouter les abandons en cours d'enquête. Le taux d'abandon risque d'être plus élevé que dans les autres enquêtes du fait de la fatigue (ou du déplaisir) que ressentent les ménages lorsque l'enquêteur leur rend des visites répétées et entreprend de leur poser des questions détaillées sur leurs revenus et leurs dépenses.

19. On ne dispose pas de bonnes études comparatives des niveaux de non-réponse dans les enquêtes sur le budget des ménages dans les pays en développement. Les enquêtes LSMS ont des taux de non-réponse de moins de 20 % (Deaton et Grosh, 2000), ce qui est considérablement plus faible que les taux enregistrés dans les enquêtes sur le budget des ménages en Europe occidentale, où ces taux peuvent atteindre 40 à 50 %. Les taux de non-réponse peuvent varier très sensiblement entre pays en développement. Dans les pays où s'exerce un puissant contrôle administratif à l'échelon communautaire, ce taux a des chances d'être faible.

C. Conception de l'échantillon

20. Les exigences qui pèsent sur la conception de l'échantillon d'une enquête sur le budget ne s'écartent guère de celles auxquelles sont soumis les autres types d'enquêtes sur les ménages. En général, on emploie un échantillon à phases multiples, et les unités primaires d'échantillonnage (UPE) sont les zones d'énumération (ZE) du recensement ou des subdivisions administratives telles que communes, villages ou quartiers. Dans la présente section, nous examinerons certaines questions propres à la conception des échantillons des enquêtes sur le budget.

1. Stratification, allocation de l'échantillon aux strates

21. Pour la stratification des UPE, on utilise généralement les régions administratives (provinces, etc.) et, à l'intérieur de ces régions, les zones urbaines et rurales. Pour les enquêtes sur le budget des ménages, une nouvelle stratification par niveau de revenu permet d'accroître l'efficacité. Dans les agglomérations et les grandes villes, on peut généralement identifier deux ou trois strates de revenu et effectuer une classifi-

cation grossière des UPE par strate (par exemple, zones à revenu élevé, intermédiaire et faible).

22. De nombreux utilisateurs ont recours aux enquêtes sur le budget des ménages et ont différentes attentes des résultats de ces enquêtes. Cela est encore plus le cas avec les enquêtes à buts multiples sur les ménages, dont font partie les enquêtes sur le budget des ménages. Le responsable de la planification de l'enquête doit souvent jongler avec des demandes contradictoires émanant d'importants utilisateurs. La comptabilité nationale [*national accounts* (NA)] est l'une des utilisations importantes des données de ces enquêtes. La NA nécessite avant tout des estimations nationales fiables des totaux pour ses comptes. Cela implique une conception de l'échantillon telle que l'échantillon soit réparti de façon homogène parmi la population (échantillon auto-pondéré) ou une conception avec suréchantillonnage des ménages à revenu intermédiaire et élevé, parmi lesquels l'activité économique est plus intense.

23. Les autres gros utilisateurs sont les responsables de la planification et de l'analyse des politiques, qui utilisent les données pour la planification, le suivi de la condition sociale et l'analyse de la pauvreté. Pour ces usages, il faut des estimations fiables pour les différentes régions et les différents groupes de population du pays de préférence à de bonnes estimations nationales. L'enquête doit porter sur un nombre suffisant de ménages de chaque région et de chaque important groupe de population (par exemple, des ménages vivant dans des villages pauvres ou reculés). Cela implique une conception de l'échantillon qui répartisse l'échantillon de façon plus ou moins égale entre les régions et, si possible, qui fasse une place suffisante aux importants groupes de population.

24. Les demandes conflictuelles évoquées précédemment doivent faire l'objet d'une certaine forme de compromis. La solution utilisée parfois en pareil cas est l'allocation selon la racine carrée, où l'échantillon est alloué aux différentes strates (régions) en proportion de la racine carrée de la taille de la strate (mesurée à son nombre de ménages). L'allocation selon la racine carrée a été utilisée pour l'Enquête sur le niveau de vie des ménages au Viet Nam et pour les enquêtes sur les ménages en Afrique du Sud.

2. Taille de l'échantillon

25. La taille des échantillons utilisés dans les enquêtes sur le budget des ménages varie selon le pays. Nombre d'enquêtes ont des échantillons de 3 000 à 10 000 ménages, mais dans les grands pays, l'échantillon peut être beaucoup plus grand. Les responsables locaux peuvent vouloir que les résultats soient détaillés au niveau géographique, au point de compromettre la qualité des données de l'enquête. Un échantillon de grande taille peut priver des ressources nécessaires les efforts tout aussi importants visant à maintenir les erreurs autres que les erreurs d'échantillonnage à des niveaux acceptables. La difficulté est de trouver le juste équilibre entre les demandes des administrations sous-nationales et les exigences budgétaires qu'implique le besoin de maintenir la taille de l'échantillon et les erreurs autres que les erreurs d'échantillonnage à des niveaux tolérables. Souvent, le responsable de la conception de l'enquête se voit contraint d'expliquer aux utilisateurs pourquoi il importe de maintenir l'équilibre entre erreurs d'échantillonnage et autres erreurs.

3. L'échantillonnage et le choix du moment

26. Les profils de dépenses et de revenu de vastes groupes de population peuvent varier considérablement selon les saisons. Il est préférable que l'enquête s'étale

sur plusieurs saisons et utilise un échantillon adéquat. Il convient de tenir particulièrement compte des longues périodes de vacances pendant lesquelles les profils de consommation varient souvent de façon considérable par rapport à ceux du reste de l'année.

27. L'un des moyens de tenir compte du problème saisonnier est d'utiliser une période de référence d'un an. Comme nous l'avons vu pour la plupart des articles et, en tout cas, pour les produits alimentaires, cela n'est pas une solution acceptable. Il est préférable de recourir à d'autres approches, par exemple :

- À des visites répétées aux mêmes ménages (avec périodes de référence répétées) étalées sur l'année, notamment sur toutes les saisons;
- À l'étude d'un ménage sur une période, par exemple, un mois (éventuellement avec plusieurs visites à ce ménage durant cette période). Les visites aux différents ménages sont étalées sur l'année selon un plan d'échantillonnage qui assure une représentation suffisante en toutes saisons. La conception doit assurer qu'en ajoutant les données mensuelles obtenues à partir d'un échantillon transversal (et en multipliant le résultat par 12), on peut reconstituer statistiquement l'année.

28. La seconde approche offre probablement la solution la plus commune au problème des saisons. Elle est utilisée dans les enquêtes sur les dépenses, notamment celles effectuées en République démocratique populaire lao, en Namibie et au Lesotho.

29. La première approche a été utilisée notamment pour l'Enquête de 1995/96 sur le revenu, la consommation et les dépenses des ménages en Éthiopie, où les ménages ont reçu deux visites, durant deux saisons différentes, et ont été interrogés sur leur revenu, leur consommation et leurs dépenses du mois précédent.

30. La seconde approche permet de tenir compte des variations saisonnières mais seulement au niveau des agrégats. Certains agrégats tels que les moyennes et les totaux annuels du revenu et des dépenses des ménages sont correctement estimés. En revanche, les mesures de la dispersion se trouvent faussées. Les totaux mensuels des différents ménages que l'on annualise en les multipliant par 12 comportent des distorsions saisonnières (dues au fait que l'enquête n'a porté que sur un mois) et des distorsions aléatoires non saisonnières (dues au fait que le même ménage a des revenus et des dépenses qui diffèrent d'un mois à un autre en raison de variations saisonnières). Ces écarts saisonniers et non saisonniers résultant de l'annualisation des totaux mensuels accentuent la distorsion par rapport à ce qu'auraient produit des totaux annuels. Les estimations de la dispersion dans les totaux annuels se trouvent donc faussées si nous utilisons la mesure de la dispersion observée dans les totaux mensuels comme estimations. On peut estimer la variation saisonnière à partir des données et l'utiliser pour réduire la distorsion. En revanche, on ne peut réduire la distorsion due à la variation au sein d'un même ménage d'un mois à un autre car on ne dispose de données pour chaque ménage que sur un mois.

31. Pour l'analyste qui veut étudier la répartition annuelle des dépenses à travers les ménages, la conception fondée sur une enquête d'un mois pose des problèmes en raison de la distorsion dans les mesures ordinaires de la dispersion (notamment, de l'écart type). Ces problèmes influent notamment sur l'analyse de la pauvreté, pour laquelle les ménages sont définis comme vivant en dessous ou au-dessus du seuil de pauvreté et leurs caractéristiques sont étudiées pour chacun de ces deux groupes. Si l'on n'apporte pas de correction, l'ampleur du phénomène de pauvreté se trouve amplifiée si moins de la moitié de la population est pauvre, et sous-estimée dans le cas contraire. Scott montre par un calcul d'un modèle que l'écart type sur les dépenses annuelles se

trouve surestimé de 36 % dans une enquête pour laquelle il n'est recueilli de données auprès des ménages que sur un mois (Scott, 1992).

D. Étude de cas : l'Enquête sur les dépenses et la consommation en République démocratique populaire lao, 1997/98

32. La République démocratique populaire lao a effectué deux enquêtes sur les dépenses et la consommation au cours de la dernière décennie. La première (LECS-1) a eu lieu en 1992/93, et la seconde (LECS-2) en 1997/98 (Comité national de planification, Centre national de statistiques de la République démocratique populaire lao, 1999). Une troisième enquête (LECS-3) est en cours.

1. Conditions générales d'exécution de l'enquête

33. D'après son dernier recensement (1995), la République démocratique populaire lao a une population de 4,5 millions d'habitants. En superficie, elle est légèrement plus vaste que la Grande-Bretagne. Les régions Nord et Est sont montagneuses. Les transports sont difficiles dans de nombreuses régions du pays : d'après le recensement de 1995, 57 % des familles rurales vivaient dans des villages sans route d'accès. La République démocratique populaire lao reste une société à prédominance rurale et agricole. L'immense majorité de la population est composée de travailleurs agricoles indépendants. Le taux d'alphabétisme des adultes est d'environ 60 %. Bien qu'elle compte plusieurs langues, la République démocratique populaire lao a une langue officielle, le lao, comprise de la majorité de sa population. Les villages sont des unités administratives bien définies, et chaque village est subdivisé en groupes de 10 à 15 ménages. D'après une évaluation assez approximative (et subjective) des conditions du travail de terrain, par comparaison avec la moyenne des pays en développement, la République démocratique populaire lao présente de plus grandes difficultés d'accès pour les enquêteurs dans les régions rurales, mais une fois atteints, les ménages sont plus enclins à coopérer.

2. Thèmes traités par l'enquête, les questionnaires

34. De vastes éléments des deux mesures macroéconomiques, « la valeur ajoutée » et « l'apport de la main-d'œuvre dans la production » ont trait à la production des ménages dans l'agriculture et aux activités informelles des ménages. Pour saisir les données sur la production des ménages, trois nouveaux modules ont été introduits dans la deuxième LECS : *a*) la tenue d'un bref journal sur le temps de travail, utilisé pour saisir l'emploi du temps d'un membre du ménage et permettre de mesurer l'apport de main-d'œuvre en heures dans l'économie lao; et *b*) deux modules sur les activités agricoles et commerciales des ménages. Cela a permis de calculer la valeur ajoutée de la production agricole et commerciale informelle des ménages.

35. Un module général sur la composition des ménages, l'éducation, l'emploi, la fécondité et la nutrition des enfants a été utilisé lors de la première entrevue. Un module « journal » a permis de couvrir toutes les transactions effectuées par le ménage au cours d'un mois. Le logement, l'accès aux biens durables, à la terre et au gros bétail ont été traités lors de la deuxième entrevue. Les questions sur le logement ont servi de base à l'imputation de valeurs au loyer. À la fin du mois, le ménage a été interrogé sur ses achats de biens durables au cours des 12 mois précédents. Un questionnaire sur le village a été présenté au chef du village. Ce questionnaire portait sur les routes et les transports, l'eau, l'électricité, les équipements de santé, les marchés locaux, les écoles, etc.

3. Méthodes de mesure

36. Le fait que la méthode du journal avait été utilisée pour la première LECS pour mesurer les transactions des ménages militait en faveur de son utilisation pour les LECS suivantes, à supposer bien entendu qu'elle ait donné de bons résultats. En changeant de méthode, on risquait de compromettre les possibilités de comparaison entre les enquêtes. La méthode du journal s'était révélée efficace lors de la LECS-1, mais seulement au prix d'un solide soutien apporté aux ménages par les enquêteurs. Quantité de ménages ne pouvaient (ou ne voulaient) tenir convenablement ce journal sans une collaboration étroite et fréquente de la part de l'enquêteur. Dans ces conditions, la méthode du journal semble être une option moins souhaitable. Cependant, il nous faut tenir compte du fait que beaucoup de villages de la République démocratique populaire lao sont difficiles d'accès. Une fois que l'enquêteur se trouve dans un village, il est souvent préférable de l'y maintenir pour les trois entrevues nécessaires avec chaque ménage, plutôt que de l'obliger à se rendre plusieurs fois dans le même village. De plus, les enquêteurs pourraient ainsi avoir des contacts fréquents avec les ménages pendant leur séjour dans le village. Le Centre national de statistiques a finalement opté pour la « méthode du journal avec l'appui de l'enquêteur » pour la LECS-2. Les enquêteurs resteraient ainsi dans le village pendant tout un mois et apporteraient aux ménages toute l'aide nécessaire à la tenue du journal.

37. Une procédure spéciale a été employée pour mesurer la consommation quotidienne de riz. On a mesuré la consommation de riz de chaque membre de la famille pendant une journée afin d'obtenir une mesure précise de la ration de riz absorbée par chacun à chaque repas. On montrait à la personne interrogée un prospectus avec des images représentant six assiettes contenant différentes quantités de riz (une boule, deux boules, etc.) et on lui demandait d'indiquer laquelle représentait sa consommation.

38. Pendant le mois, on a choisi une période de 24 heures pour enregistrer l'emploi du temps de la famille. Ce journal indiquant l'emploi du temps utilisé pour la LECS-2 avait été mis au point conjointement par Statistics Sweden et le Centre de recherche sur le changement microsocial du Conseil de la recherche économique et sociale de l'Université de l'Essex. On avait cherché essentiellement à l'alléger, afin qu'il puisse être utilisé conjointement avec d'autres instruments sans surcharger les répondants. Il a été demandé à un seul membre de la famille (choisi au hasard parmi les membres âgés de 10 ans ou plus) de remplir ce journal pour une journée désignée à cet effet. L'enquêteur choisissait les répondants au hasard, de telle manière que le nombre choisi chaque jour de la semaine devait rester constant.

39. Le journal sur l'emploi du temps contenait 22 activités prédéfinies et mettait l'accent sur les activités économiques. Pour certaines de ces activités, l'enquêteur cherchait à recueillir des informations supplémentaires au moment où il ramassait ces journaux. Ceux qui répondaient « ai travaillé comme employé » se voyaient demander s'ils avaient travaillé comme ouvrier agricole, dans le secteur public, dans le secteur privé ou ailleurs. À ceux qui répondaient « à mon compte », on demandait quel rôle ils avaient joué dans leur entreprise. Les réponses étaient classées selon une liste d'environ 50 catégories fondées sur la Classification internationale type, par industrie, de toutes les branches d'activité économique (CITI) et le Système de comptabilité nationale, 1993 [Commission des Communautés européennes, Fonds monétaire international, Organisation de coopération et de développement économiques, Nations Unies et Banque mondiale (1993)].

4. Conception de l'échantillon, travail de terrain

40. Les zones d'énumération (ZE) du recensement servent d'unités primaires d'échantillonnage (UPE). Les UPE ont été stratifiées en 18 provinces et, à l'intérieur de chaque province, entre zones rurales et urbaines. Les ZE rurales ont été en outre stratifiées en ZE « avec accès à une route » et « sans accès à une route ». Un échantillon de 25 UPE a été alloué à chaque province. Une allocation complémentaire par zone urbaine ou rurale a été effectuée, et la partie urbaine s'est vu attribuer une fraction d'échantillonnage de 50 % plus grande que la partie rurale. Les UPE ont été sélectionnées selon une probabilité systématique proportionnelle à la taille [*systematic probability proportional to size* (PPS)] dans chaque province, produisant un échantillon de 450 UPE.

41. La liste des ménages des UPE sélectionnées a été dressée avant l'enquête, et 20 ménages ont été choisis selon un échantillonnage systématique dans chaque UPE, produisant un échantillon total de 9 000 ménages. L'échantillonnage dans le temps s'est effectué par attribution aléatoire de l'échantillon provincial sur la période de 12 mois, donnant deux (et, dans un cas, trois) villages par mois.

42. Il fallait une équipe de deux enquêteurs pour le travail dans le village. Les enquêteurs ont été sélectionnés parmi le personnel permanent des bureaux provinciaux de statistiques. Beaucoup avaient participé à la première LECS. La formation s'est déroulée sur une période de deux semaines.

E. Expériences, leçons apprises

1. Méthodes de mesure, non-réponse

43. Les enquêteurs ont passé beaucoup de temps auprès des ménages à aider les répondants à enregistrer toutes les transactions relatives au ménage proprement dit ainsi qu'à ses activités commerciales et agricoles. On a des raisons de croire que ce travail laborieux et long a contribué à améliorer la qualité des réponses. Certaines anecdotes semblent indiquer que, souvent, les fréquentes visites de l'enquêteur au ménage ont permis de créer entre eux un climat de confiance. Elles ont également donné aux enquêteurs amplement le temps d'analyser les liens compliqués entre la consommation du ménage et sa production agricole ou ses activités commerciales.

44. Quelques contrôles de qualité ont été effectués. Les estimations de la consommation de riz ont été vérifiées en regard des données sur la production agricole extérieure et trouvées raisonnablement compatibles. Il a également été procédé à une vérification des niveaux de consommation entre la première et la deuxième période de deux semaines de tenue du journal. Il n'y avait aucune indication d'un fléchissement de cette consommation durant la seconde période, et il y avait de légères différences entre les deux périodes dans le nombre d'entrées dans le journal. Par ailleurs, les estimations de la consommation totale étaient comparables entre les deux périodes.

45. Le fait qu'il y avait de légères différences dans la consommation au niveau des agrégats entre la première et la seconde période de deux semaines de tenue du journal amène à se poser la question de savoir si des périodes plus courtes auraient suffi pour enregistrer la consommation.

46. Le taux de non-réponse était faible, de 3,1 % seulement. Il était très faible dans les zones urbaines (0,6 % seulement) et plus élevé, quoique toujours faible, dans les zones rurales (3,9 %). Ce taux était quelque peu sous-estimé. On a utilisé la substitution pour les cas de non-réponse, mais les activités de notification des résultats par les enquêteurs laissaient à désirer, de sorte qu'il est difficile d'évaluer le niveau exact de non-réponse et de différencier entre l'impossibilité d'établir le contact avec les mé-

nages et les cas de refus de répondre. Le nombre de refus était très faible. Toutes les expériences des enquêtes sur les ménages en République démocratique populaire lao indiquent que les ménages se sentent obligés de participer aux enquêtes des pouvoirs publics. De plus, le chef de village leur donne pour instructions de participer.

2. Conception de l'échantillon, erreurs d'échantillonnage

47. Il ressort de l'analyse de la variance et des structures de coût que la taille optimale des échantillons à l'intérieur des UPE (zones d'énumération) se situe entre 8 et 12 ménages. Autrement dit, le nombre de 20 ménages utilisé comme échantillon dans cette enquête était supérieur au niveau optimal (Pettersson, 2001).

48. Les calculs montrent également que l'allocation de l'échantillon à égalité entre les provinces entraîne des erreurs d'échantillonnage dans les estimations nationales d'environ 20 % plus élevées qu'une allocation proportionnelle. Les coefficients de variation (CV) étaient généralement inférieurs à 5 % pour les estimations au niveau national. L'échantillon était plus petit dans les zones urbaines que dans les zones rurales (2 008 ménages contre 6 874), mais les CV étaient comparables entre estimations urbaines et rurales, ce qui est dû en partie à des effets de conception moins prononcés dans les zones urbaines.

49. Les effets de conception étaient relativement élevés dans les zones rurales, considérablement plus élevés que dans les zones urbaines (voir tableau XXIV.1). Cela s'explique par le fait que les villages ruraux sont socio-économiquement homogènes. Comme la plupart des UPE rurales sont constituées d'un village, les UPE sont elles aussi homogènes. Dans les petites et grandes villes, il y a relativement peu de regroupement en quartiers riches et pauvres : les familles riches vivent aux côtés des familles pauvres dans tous les quartiers de la ville. C'est pourquoi nombre d'UPE urbaines contiennent à la fois des familles riches et des familles pauvres, ce qui rend les UPE urbaines relativement hétérogènes.

Tableau XXIV.1
Effets de conception sur la consommation des ménages
et sur la possession de biens durables

	National	Urbain	Rural
Consommation mensuelle totale par ménage	5,4	3,8	7,7
Consommation alimentaire mensuelle par ménage	5,8	4,4	6,8
Proportion de ménages possédant une automobile	2,1	1,3	3,3
Proportion de ménages possédant un téléviseur	5,4	3,1	6,8
Proportion de ménages possédant une radio	4,5	2,7	4,8
Proportion de ménages possédant un magnétoscope	5,5	3,9	6,1

50. Chaque ménage de l'échantillon a été étudié pendant un mois, et l'échantillon a été étalé de façon homogène sur une période de 12 mois. Cela a posé des problèmes lorsque l'enquête a été utilisée pour estimer les taux de pauvreté (voir section C.3). Les variations saisonnières ont été estimées à partir des données et prises en compte dans les estimations de l'année. En revanche, il n'a pas été possible d'estimer la variation aléatoire non saisonnière entre ménages d'un mois à un autre. De ce fait, la dispersion de la consommation des ménages a été légèrement surestimée, de même que les taux de pauvreté.

3. Expériences tirées de l'utilisation du journal sur l'emploi du temps

51. On ignore le nombre de journaux remplis par les répondants et le nombre de ceux remplis par l'enquêteur. Certains signes indiquent toutefois qu'en règle générale, les enquêteurs ont apporté une aide considérable à la plupart des répondants, encore que l'on note certaines différences régionales.

52. La sélection au hasard d'une personne par famille n'a pas donné de bons résultats. Le calcul de la répartition par âge et par sexe n'a guère réussi à faire respecter les règles de la sélection au hasard. Il semble que, dans bien des cas, l'enquêteur n'a pas vraiment appliqué ce mode de sélection et a permis des substitutions, probablement pour des raisons pratiques. D'après les calculs, les hommes d'âge actif (de 15 à 64 ans) étaient sur-représentés et les jeunes des deux sexes (de 10 à 14 ans) ainsi que les personnes âgées (de 65 ans et plus), surtout les femmes, étaient sous-représentées dans la sélection (Johansson, 2000) [voir tableau XXIV.2]. Il faut modifier la procédure employée pour obtenir une meilleure représentativité des données sur l'emploi du temps. Si le module sur l'emploi du temps est conçu pour refléter principalement les activités économiques, les éléments les plus jeunes et les plus âgés de la population n'ont pas besoin d'être inclus. En revanche, il convient d'inclure ces catégories de population pour les programmes sociaux concernant plus particulièrement le travail des enfants et la situation des personnes âgées.

Tableau XXIV.2

Ratio entre le nombre effectif et le nombre attendu de personnes dans l'échantillon pour le journal sur l'emploi du temps

Âge	Ratio nombre effectif/nombre attendu		
	Hommes	Femmes	Total
10-14 ans	0,41	0,49	0,45
15-64 ans	1,33	1,04	1,18
65 ans et plus	0,59	0,29	0,43
Total	1,11	0,90	1,00

4. Utilisation de LECS-2 pour les estimations du PIB

53. Les expériences de l'utilisation de modules de mesure de la valeur de la production des ménages et du coût de leurs apports, ainsi que du temps utilisé, ont été encourageantes. Cette utilisation a permis de renforcer considérablement la base statistique des estimations du produit intérieur brut (PIB). L'enquête fournit à présent des données importantes pour la comptabilité nationale en ce qui concerne : a) la valeur ajoutée de la production du ménage; b) l'apport de main-d'œuvre à l'économie; et c) le niveau et la structure de la consommation privée.

54. Dans la nouvelle estimation de base du PIB pour 1997, la production des ménages dans les domaines de l'agriculture et des activités économiques informelles représentait 64 % du PIB et l'utilisation, une proportion encore plus élevée du PIB. Environ 80 % de l'apport de main-d'œuvre à l'économie venaient de la production des ménages, tant en agriculture que pour ce qui est des activités économiques informelles (Johansson, 2000).

F. Conclusion

55. Ce chapitre a porté sur les questions concernant la conception d'enquêtes ayant pour but de mesurer « le budget des ménages ». Il s'est concentré sur les enquêtes portant sur les estimations de la consommation totale des ménages et de leur production et où ces estimations servent à leur tour d'élément pour la comptabilité nationale et les statistiques économiques nationales en général. Pour un examen plus approfondi des questions de conception, le lecteur est invité à se reporter à d'autres publications [par exemple, Deaton et Grosh (2000) et Nations Unies (1989)].

56. L'étude de cas utilisée dans ce chapitre est quelque peu inhabituelle du point de vue du temps passé par l'enquêteur avec chaque ménage. Des considérations de précision et les conditions du travail de terrain militaient en faveur de cette conception pour l'enquête effectuée en République démocratique populaire lao. L'utilisation de la méthode du journal parmi une population à faible taux d'alphabétisme a obligé à recourir à un appui quasi quotidien envers de nombreux ménages. La méthode de la tenue d'un journal avec l'appui de l'enquêteur a été jugée nécessaire pour enregistrer avec précision la consommation des ménages lao. D'autres méthodes moins coûteuses pourraient donner des estimations d'une qualité acceptable dans d'autres pays.

RÉFÉRENCES

- Banque mondiale (1992). *Indonesia: public expenditures, prices and the poor*. Indonesia Resident Mission 11293-IND, Djakarta. Cité dans Deaton (1997).
- Commission des Communautés européennes, Fonds monétaire international, Organisation de coopération et de développement économiques, Nations Unies et Banque mondiale (1993). *Système de comptabilité nationale, 1993*, numéro de vente : E.94.XVII.4.
- Deaton, A. (1997). *The Analysis of Household Surveys, A Micro Econometric Approach to Development Policy*. Baltimore, Maryland, et Londres, Johns Hopkins University Press.
- _____, et M. Grosh (2000). Consumption. Dans *Designing Household Survey Questionnaires in Developing Countries: Lessons from 15 Years of Living Standards Measurement Study*, M. Grosh et P. Glewwe, eds. Washington, D.C., Banque mondiale.
- Glewwe, P. et Yansaneh, I. (2001). *Recommendations for Multi-Purpose Household Surveys from 2002 to 2010*. Report of Mission to the General Statistics Office, Viet Nam.
- Johansson, S. (2000). *A Household Survey Program for Lao PDR. Report on a Short-Term Mission to Vientiane, August 7-21, 2000*. Stockholm, International Consulting Office, Statistics Sweden.
- Joliffe, D., et K. Scott (1995). The sensitivity of measures of household consumption to survey design: results from an experiment in El Salvador. Washington, D.C., Policy Research Department, World Bank.
- Mahalanobis, P. C., et B. Sen (1954). On some aspects of the Indian National Sample Survey. *Bulletin of the International Statistical Institute*, vol. 34.
- Nations Unies (1989). *National Household Survey Capability Programme: Household Income and Expenditure Surveys: A Technical Study*. DP/UN/INT-88-X01/6E. Department of Technical Co-operation for Development and Statistical Office.
- Pettersson, H. (2001). *Sample Design for the Household Surveys: Report from a Mission to the National Statistics Centre, Lao P.D.R., 19 février- 2 mars 2001*, Stockholm, International Consulting Office. Statistics Sweden.

- Rydenstam, K. (2000). The « light » time diary approach: report on some Lao PDR and Swedish actions and experiences. Paper prepared for the United Nations Expert Group Meeting on Methods for Conducting Time-Use Surveys, 23-27 octobre 2001.
- Scott, C. (1992). Estimation of annual expenditure from one-month cross-sectional data in a household survey. *Inter-Stat Bulletin*, vol. 8, pp. 57-65.
- _____, et B. Amenuvegbe (1990). *Effect of Recall Duration on Reporting of Household Expenditures: An Experimental Study in Ghana*. Social Dimensions of Adjustment in Sub-Saharan Africa Working Paper, No. 6. Washington, D.C., Banque mondiale.
- State Planning Committee, National Statistical Centre of Lao People's Democratic Republic (1999), *The household of Lao PDR: Social and economic indicators: Lao Expenditure and Consumption Survey 1997/98*. Vientiane.
- Statistical Institute and Planning Institute of Jamaica (1996). *Jamaica Survey of Living Conditions 1994*. Kingston.

Chapitre XXV

Les enquêtes sur les ménages dans les pays en transition

JAN KORDOS

École d'économie de Varsovie,
Office central de la statistique
Varsovie, Pologne

RÉSUMÉ

Le présent chapitre offre un aperçu des principaux aspects de la conception et de l'exécution d'enquêtes par sondage sur les ménages dans les pays en transition au cours de la dernière décennie, 1991-2000. Il donne également des informations tirées de 14 pays en transition sur certains aspects opérationnels de ces enquêtes. Les bureaux de statistiques de ces pays ont fourni ces informations en 2001 en remplissant des questionnaires spéciaux qu'ils ont par la suite, dans certains cas, actualisés.

Ce chapitre comprend deux sections : la section A donne un examen général des enquêtes sur les ménages dans les pays en transition. La section B présente des études de cas de telles enquêtes dans certains de ces pays.

La section A présente une synthèse des principaux aspects des enquêtes sur les ménages dans les pays en transition. Il examine en particulier deux principaux types d'enquêtes : l'enquête sur le budget des ménages [*household budget survey* (HBS)] et l'enquête sur la main-d'œuvre [*labour-force survey* (LFS)]. Il analyse les éléments suivants de ces enquêtes : le cadre d'échantillonnage, la conception de l'échantillon, la taille des échantillons, l'estimation des erreurs d'échantillonnage, les taux de non-réponse, les coûts des enquêtes et les effets de conception. Les pays en transition avaient déjà une certaine expérience de l'enquête HBS, quoique leur conception ait dû être repensée dans chaque pays. L'enquête LFS est une enquête d'un type entièrement nouveau et n'est apparue dans différents pays en transition qu'au cours de la dernière décennie, parfois avec une assistance technique extérieure. La section A conclut par des recommandations pour améliorer les enquêtes sur les ménages dans les pays en transition en tenant compte des recensements de la population et des logements de 2000.

La section B présente des études de cas sur les pays suivants : Estonie, Hongrie, Lettonie, Lituanie, Pologne et Slovénie. Les descriptions soulignent les principaux aspects des enquêtes HBS, LFS et autres enquêtes sur les ménages effectuées dans chacun de ces pays.

Termes clés : enquête sur le budget des ménages, enquête sur la main-d'œuvre, coût de l'enquête, effet de conception, erreur d'échantillonnage, taux de non-réponse.

A. Examen général des enquêtes sur les ménages dans les pays en transition

1. Introduction

1. La présente section a pour but de présenter certains aspects de la conception et de l'exécution des enquêtes sur les ménages dans certains pays en transition, en par-

ticulier dans certains pays d'Europe centrale et orientale et de la Fédération de Russie au cours de la dernière décennie. Les profondes différences qui existent entre les divers types d'enquêtes par sondage sur les ménages, que ce soit sur la substance, les unités de réponse, la périodicité, la conception de l'échantillon ou les méthodes de collecte des données, se traduisent par des différences dans les niveaux de coût et les taux de non-réponse. Le présent chapitre porte sur la conception et l'exécution de deux types d'enquêtes par sondage sur les ménages, l'enquête sur le budget des ménages (HBS) et l'enquête sur la main-d'œuvre (LFS). Cependant, d'autres types d'enquêtes sur les ménages effectuées dans les pays en transition au cours de la dernière décennie sont également cités.

2. Avant d'examiner les enquêtes par sondage sur les ménages effectuées dans les pays en transition au cours de la dernière décennie, nous présenterons une description générale des enquêtes sur les ménages effectuées dans ces pays avant la période de transition afin de faire mieux comprendre l'évolution de ces enquêtes dans ces pays.

3. Pour la préparation de ce chapitre, nous avons conçu un questionnaire spécial que nous avons envoyé aux 14 pays en transition suivants : Bélarus, Bulgarie, Croatie, République tchèque, Estonie, Hongrie, Lettonie, Lituanie, Pologne, Roumanie, Fédération de Russie, Slovaquie, Slovénie et Ukraine.

4. Huit pays ont préparé des études approfondies qui ont été publiées dans *Statistics in Transition* [vol. 5, n° 4 (juin 2002)].

5. Une attention particulière a été prêtée à la conception et à l'exécution des enquêtes par sondage sur les ménages, notamment à certaines questions telles que les cadres d'échantillonnage, la conception de l'échantillon, la taille de l'échantillon, les méthodes d'estimation des paramètres et des erreurs d'échantillonnage, les taux de non-réponse, les coûts des enquêtes et de leurs composantes, les effets de conception et leur utilisation dans l'analyse statistique. Ce chapitre décrit également les plans à venir pour améliorer ces enquêtes après le cycle des Recensements de la population et des logements de 2000.

2. Enquêtes par sondage sur les ménages dans les pays d'Europe centrale et orientale et en URSS avant la période de transition (1991-2000)

6. Il n'est pas facile d'analyser objectivement les enquêtes par sondage sur les ménages en Europe centrale et orientale et en URSS avant la période de transition. On sait très bien que ces pays avaient un système de statistiques centralisé, et que les rapports exhaustifs ou les recensements étaient les principales formes de collecte de données. Cependant, certaines publications décrivent des enquêtes sur les ménages effectuées dans ces pays avant cette période, et l'on sait que des conférences, séminaires et groupes de travail étaient organisés pour discuter des méthodes utilisées pour ces enquêtes.

7. Les anciens pays communistes, à savoir les pays d'Europe centrale et orientale et l'Union soviétique, avaient un système d'enquêtes par sondage sur les ménages dont les plus importantes étaient les enquêtes sur le budget familial [*family budget surveys* (FBS)]. De vastes enquêtes étaient également lancées périodiquement sur les conditions de vie, de même que des enquêtes sur les revenus, des micro-recensements, des enquêtes sur la santé, des enquêtes sur l'emploi du temps et différents types d'enquêtes sociales et démographiques par sondage.

8. À partir des années 1950, les enquêtes sur le budget de famille ont été établies selon la méthodologie soviétique fondée sur ce qui était appelé l'approche par

secteurs (Postnikov, 1953), qui consistait à choisir des ménages parmi les employés de certaines entreprises dans chaque secteur d'activité. Les ménages choisis, qui participaient à l'enquête pendant plusieurs années, tenaient un journal de leurs revenus et de leurs dépenses. L'échantillon n'était pas renouvelé et ne comprenait que des ménages dont certains membres étaient employés dans des entreprises socialistes, excluant les personnes qui vivaient trop loin des entreprises sélectionnées. Dans chaque secteur, les ménages étaient sélectionnés selon un processus à deux phases. À la première phase de la sélection, un nombre déterminé d'entreprises (ou d'autres lieux de travail) du pays étaient sélectionnées selon une probabilité proportionnelle au nombre total de leurs employés. À la seconde phase, dans chaque entreprise sélectionnée, le même nombre de ménages était systématiquement sélectionné à partir d'une liste d'employées stratifiée par type de groupe économique. Ces groupes étaient d'abord classés selon l'ampleur des salaires. À chaque phase de la sélection, les unités d'échantillonnage étaient choisies systématiquement à partir du milieu de « l'intervalle d'échantillonnage ». Il était supposé que cette méthode de sélection de l'échantillon était autopondérée pour chaque secteur. Après la sélection, une procédure spéciale était appliquée pour vérifier la représentativité de l'échantillon à l'aide de données sur la moyenne des salaires. Au début, les taux de non-réponse étaient faibles, et il y avait des différences notables entre les pays.

9. Durant les années 1959-1962, une attention particulière a été accordée à l'amélioration et à l'unification des FSB. Pour ce faire, un Comité permanent de statistiques du Conseil d'assistance économique mutuelle (CAEM) a été créé comme groupe de travail spécial dont les membres ont été choisis parmi les pays d'Europe centrale et orientale et l'Union soviétique. Certains progrès ont été notés dans des domaines méthodologiques tels que la définition des concepts, les classifications et la conception des questionnaires. Certains pays ont exprimé des doutes sur l'approche par secteur, soulignant les inconvénients que présente le fait de maintenir le même ménage dans l'enquête pendant plusieurs années. Dans certains pays, les taux de non-réponse ayant augmenté régulièrement, il a été suggéré d'adopter une méthode de rotation pour la sélection de l'échantillon et de raccourcir la durée de participation du même ménage à l'enquête. Dans les années 1960, certains pays ont essayé une « approche territoriale », selon laquelle les ménages étaient sélectionnés à partir des zones d'énumération du recensement et les logements étaient stratifiés par région. Le renouvellement des ménages de l'échantillon a permis de raccourcir la durée de participation des ménages à l'enquête [Główny Urząd Statystyczny (GUS), 1971a; Kordos, 1985, 1996]¹. Dans certains pays d'Europe centrale et orientale (Bulgarie, Tchécoslovaquie, Hongrie, Pologne, Roumanie), la méthodologie employée pour les enquêtes HBS a commencé à changer.

¹ GUS = Office central de la statistique de Pologne.

10. Après quelques expériences, en 1971, la Pologne a adopté l'approche territoriale pour l'enquête HBS en 1971 et, en 1982, elle a appliqué la méthode par rotation [Główny Urząd Statystyczny (GUS), 1971a; Kordos, 1982, 1985; Lednicki, 1982].

11. En Hongrie, après la mise en place d'un Système unifié d'enquêtes sur les ménages au milieu des années 1970, l'enquête sur le budget des ménages est devenue un processus continu qui a duré pendant toute la période 1976-1982; de 1983 à 1991, l'enquête a été effectuée tous les deux ans; et depuis 1993, elle est à nouveau continue. Les enquêtes sur les revenus, lancées en 1963, ont été effectuées deux fois tous les dix ans. Un certain nombre d'enquêtes sur les ménages ont été organisées dans le cadre du Système unifié, en particulier dans les années 1980, notamment une enquête sur l'emploi du temps, une enquête sur le prestige des diverses professions, une enquête sur les conditions de vie et la stratification sociale, etc. (Mihalyffy, 1994; Élterő et Mihalyffy, 2002).

12. D'autres enquêtes sur les ménages ont été effectuées dans ces pays pendant la période qui a précédé la transition. Le Comité permanent de statistiques du CAEM a inclus dans son plan de travail 1968-1970 un thème consacré aux « Possibilités d'application plus large des méthodes d'échantillonnage dans les enquêtes statistiques des pays membres du Conseil d'assistance économique mutuelle ». En avril 1970, la Pologne a été chargée d'organiser un séminaire sur ce thème et de préparer le document principal (Kordos, 1970). Neuf pays (la Bulgarie, la Hongrie, la Mongolie, la Pologne, la République démocratique allemande, la Roumanie, la Tchécoslovaquie, l'Union soviétique et la Yougoslavie) ont participé à ce séminaire. Chaque pays a présenté un document en russe, et ces documents ont été publiés par la suite en polonais dans un volume spécial [Główny Urząd Statystyczny (GUS), 1971a]. Des études méthodologiques ont été également présentées et publiées dans un second volume [Główny Urząd Statystyczny (GUS), 1971b]. Ces études permettent de se faire une idée générale du type d'enquêtes sur les ménages qui avaient été effectuées dans ces pays jusqu'en 1970.

13. Par ailleurs, plusieurs conférences internationales ont été consacrées aux enquêtes sur les ménages, et particulièrement sur le budget des ménages. Les statisticiens polonais ayant participé à ces réunions internationales ont préparé des rapports détaillés qui ont été publiés dans des journaux statistiques polonais. L'auteur a participé à un séminaire européen de statistiques consacré aux enquêtes sur les ménages qui s'est tenu à Vienne en 1961 (Kordos, 1963) et à la deuxième conférence internationale sur la méthodologie des enquêtes sur les ménages qui s'est tenue à Genève en 1981 (Kordos, 1981).

14. Les enquêtes sur les ménages effectuées dans ces pays ont été examinées à la Conférence internationale sur les statistiques économiques pour les économies en transition : l'Europe de l'Est dans les années 1990, tenue à Washington, D.C., du 14 au 16 février 1991 (Garner *et al.*, 1993).

15. À partir de ces publications, on peut constater que les méthodes d'échantillonnage étaient également utilisées pour accélérer le traitement des données des recensements de la population et des logements (Bulgarie, Pologne, République démocratique allemande, Tchécoslovaquie et Yougoslavie); pour effectuer des microrecensements (Hongrie, Pologne, Tchécoslovaquie, URSS et Yougoslavie); pour étudier les conditions de vie (Bulgarie, Hongrie, Pologne, Roumanie, URSS); pour des enquêtes post-énumération après les recensements de la population et des logements (Bulgarie, Hongrie, Mongolie, Pologne, Roumanie, Tchécoslovaquie, URSS, Yougoslavie); et pour étudier l'emploi du temps (Bulgarie, Hongrie, Pologne, Roumanie). Il y avait de grandes différences de développement statistique entre ces pays, ce qui a eu un certain impact sur le progrès des enquêtes sur les ménages dans la période de transition.

3. Les enquêtes sur les ménages dans la période de transition

16. La présente section porte sur la méthodologie et l'exécution des enquêtes par sondage sur les ménages effectuées dans la période de transition, c'est-à-dire entre 1991 et 2000. Ces enquêtes étaient extrêmement longues et se sont modifiées durant cette période par rapport à celles d'avant 1990. Les enquêtes sur le budget des ménages se sont améliorées et continuent de s'améliorer, et pour la première fois dans chaque pays, une nouvelle enquête, sur la main-d'œuvre, a été introduite ou est sur le point de l'être. En outre, de nouvelles enquêtes par sondages sur les ménages – sur le bien-être et la santé de la population ou sur les conditions de vie, ainsi que d'autres enquêtes démographiques et sociales – ont été lancées.

Tableau XXV.1

Nouvelles enquêtes sur le budget des ménages et sur la main-d'œuvre dans les pays en transition, 1992-2000 : première année, périodicité et année de la dernière refonte

Pays	Première année		Périodicité		Année de la dernière refonte	
	HBS	LFS	HBS	LFS	HBS	LFS
Bélarus	1995	-	Trim.	-	1995	-
Bulgarie	1992	1993	Cont.	Trim.	2000	2001
Croatie	1998	1996	Cont.	Deux fois	2000	2000
République tchèque	1991	1993	Cont.	Cont. ^a	1999	2000
Estonie	1995	1995	Cont.	Cont.	1999	2000
Hongrie	1976	1996	Cont.	Cont.	1997	1997
Lettonie	1995	1995	Cont.	Deux fois	1998	1999
Lituanie	1992	1994	Cont.	Deux fois	1996	1997
Pologne	1982	1992	Cont.	Cont. ^b	2000 ^c	1999
Roumanie	2001	1994	Cont.	Cont. ^d	2000	2001
Fédération de Russie	1997	1992	Cont. ^e	Trim.	1996	1998
Slovaquie	2003	1993	Cont. ^f	Cont.	2002	1999
Slovénie	1999	1993	Cont.	Cont.	1997	Annuelle ^g
Ukraine	1999	1999	Cont.	Trim.	2000	1999

Source : Données tirées des questionnaires soumis par les pays.

Note : Trim. signifie trimestriel; Cont. signifie continue; Deux fois signifie deux fois par an. Un tiret (-) signifie sans objet.

^a Depuis 2000.

^b Depuis le quatrième trimestre de 1999.

^c Depuis 1982, modifiée trois fois.

^d Depuis 1996.

^e Continue depuis 1952 mais modifiée en 1996. Une nouvelle enquête a démarré en 1997.

^f Continue depuis 1957, mais modifiée en 2002. Une nouvelle enquête a démarré en 2003.

^g Questions spéciales ajoutées chaque année.

17. Nous commencerons par un examen des enquêtes HBS et LFS. D'autres enquêtes périodiques ou ponctuelles sont également décrites en termes généraux. Ensuite, nous accorderons une attention particulière à certains aspects méthodologiques communs à toutes les enquêtes sur les ménages, tels que la construction du cadre d'échantillonnage, la conception de l'échantillon, la méthode d'estimation, l'erreur d'échantillonnage, l'effet de conception, les coûts de l'enquête, les cas de non-réponse et les améliorations qu'il est prévu d'apporter à l'avenir aux enquêtes sur les ménages.

18. Durant la dernière décennie, dans presque tous les pays en transition, l'enquête sur le budget des ménages (HBS) a été refondue et une nouvelle enquête a été introduite. Comme il n'y avait pas d'enquêtes sur la main-d'œuvre (LFS) avant la période de transition, de nouvelles enquêtes de ce type ont été conçues et exécutées. Le tableau XXV.1 indique l'année de démarrage de la nouvelle enquête HBS et de la nouvelle enquête LFS, leur périodicité et la dernière année de leur refonte.

19. Comme le montre le tableau XXV.1, les nouvelles enquêtes HBS, après avoir été restructurées et adaptées aux normes de l'Office statistique des Communautés européennes (Eurostat, 1997), sont généralement devenues des enquêtes continues. Les enquêtes LFS ont été introduites dans les pays en transition au cours de la période 1992-1999.

4. Enquêtes sur le budget des ménages

20. Les enquêtes sur le budget des ménages sont une vieille tradition dans les pays en transition. Une grande attention est accordée à ces enquêtes en raison du rôle particulier qu'elles jouent dans l'analyse des conditions de vie de la population et dans le calcul des indices de prix à la consommation. Diverses méthodes ont été expérimentées et plusieurs tentatives ont été faites pour améliorer la méthodologie et l'organisation de ces enquêtes. Certains pays, tels que la Bulgarie, la Hongrie, la Pologne et la Roumanie, ont commencé à apporter des améliorations à la méthodologie de ces enquêtes dans les années 1970 et 1980. Au début des années 1990, d'autres pays ont suivi leur exemple. Les enquêtes ont été repensées et adaptées aux normes d'Eurostat (Eurostat, 1997). Eurostat est prêt à aider les États membres, ainsi que les autres pays intéressés, à améliorer leurs méthodes et procédures d'enquête en leur donnant des directives et en leur apportant un appui technique direct (Eurostat, 1995, 1996, 1998a, 1998b). C'est ainsi que de nouveaux concepts, de nouvelles définitions et de nouvelles classifications ont été adoptés et de nouveaux journaux et questionnaires ont été conçus. Pour la première fois, ces enquêtes sont également utilisées comme source d'informations dans la tenue des comptes nationaux pour mesurer la consommation finale des ménages à un niveau d'agrégats.

21. Toutes les enquêtes HBS se limitent à la population vivant dans des cellules familiales. Les établissements collectifs ou institutionnels (hôpitaux, foyers pour personnes âgées, pensionnats, prisons, casernes, etc.) sont exclus. À l'exception de la République tchèque et de la Slovaquie, tous les 14 pays en transition ont procédé à une refonte de leurs enquêtes HBS.

5. Enquêtes sur la main-d'œuvre

22. Pour les pays en transition, les enquêtes sur la main-d'œuvre sont un nouveau concept, apparu seulement après 1992. Eurostat et des représentants des Offices nationaux de statistique et des ministères du travail, aux fins de discuter des aspects techniques de ces enquêtes, ont participé régulièrement plusieurs fois par an aux réunions du Groupe de travail sur les statistiques de l'emploi tenues à Luxembourg (Eurostat, 1998a, 1998b). Ainsi, les enquêtes LFS ont pu être exécutées selon les recommandations de l'Organisation internationale du Travail (OIT) et les méthodes et définitions d'Eurostat (Eurostat, 1998a).

23. Depuis 1989, le Bureau de statistique de l'OIT s'emploie activement à aider les pays d'Europe centrale et orientale et de l'ex-URSS à réviser et à restructurer radicalement leurs systèmes de statistiques sur la main-d'œuvre afin de répondre aux nouvelles exigences découlant de leur transition vers l'économie de marché. Cette assistance technique a été dispensée sous la forme d'un certain nombre de cours, séminaires, conférences et visites d'experts.

24. En ce qui concerne l'enquête LFS, les experts de l'OIT ont effectué des missions en Fédération de Russie (à deux reprises en 1992 pour la préparation de l'enquête LFS pilote, et en 1993 pour l'enquête LFS en vraie grandeur); en Ukraine (en novembre 1991 et novembre 1992 pour la préparation d'une enquête LFS pilote et en novembre 1993 pour l'exécution d'une enquête à titre d'essai); en Bulgarie (décembre 1991, juillet et octobre 1992, avril 1993 et février 1994); en Slovénie (octobre 1993); au Bélarus (novembre 1993 et septembre 1994 pour la préparation d'une enquête pilote et suivi); au Kazakhstan (mars et juin 1993 pour examiner la faisabilité de lancement d'une enquête LFS pilote). Par ailleurs, trois stages de formation sur le tas sur la prépa-

ration et l'exécution d'une enquête LFS ont été organisés à l'intention de spécialistes russes et ukrainiens en Norvège (1991) et en Allemagne (1991 et 1992).

25. En 1994 (31 août-2 septembre), l'OIT a organisé à Minsk une Conférence internationale sur la restructuration des statistiques de main-d'œuvre dans les pays en transition. L'objectif immédiat de cette conférence était de faire le point de ce qui avait été fait et de ce qu'il restait à faire pour produire des statistiques fiables et homogènes sur le marché du travail afin de répondre aux besoins d'information et de formulation de politiques dans les pays en transition. Tous les documents préparés pour cette conférence ont été publiés dans une édition spéciale de *Statistics in Transition* (vol, 2, n°1, mars 1995).

26. Certains aspects de la conception et de l'exécution des enquêtes LFS dans les 14 pays en transition étudiés méritent examen. Comme le montre le tableau XXV.1, treize de ces pays ont déjà entrepris des enquêtes LFS, et le Bélarus prévoit d'en lancer une prochainement. Sept de ces 14 pays (l'Estonie, la Hongrie, la Pologne, la République tchèque, la Roumanie, la Slovaquie et la Slovénie) effectuent des enquêtes continues, ce qui veut dire que les semaines de référence sont réparties de façon régulière sur l'ensemble de l'année. Dans trois pays (Bulgarie, Fédération de Russie et Ukraine), l'enquête est trimestrielle et dans trois autres (Croatie, Lituanie et Lettonie), elle est semestrielle. En Estonie, jusqu'en 1999, l'enquête a été effectuée une fois par an (au printemps), mais depuis 2000, elle a lieu tous les trimestres. Tous les pays envisagent de restructurer prochainement leurs enquêtes LFS, en utilisant comme base les résultats des recensements de la population et des logements pour améliorer leur cadre d'échantillonnage, la conception de l'échantillon et la méthode d'estimation.

6. Traits communs de la conception des échantillons et de l'exécution des enquêtes HBS et LFS

27. Les enquêtes HBS et LFS représentent des types d'enquête sur les ménages d'un genre particulièrement différent. Toutefois, dans la mesure où certains éléments de leur méthodologie et de leur exécution (tels que le cadre d'échantillonnage, la conception de l'échantillon, la méthode d'estimation, l'estimation de l'erreur d'échantillonnage, l'effet de conception, le coût, les taux de non-réponse et les plans d'amélioration à venir des enquêtes) sont communs aux deux types, il est jugé utile de les examiner ensemble.

28. Les différents pays ont suivi des procédures assez semblables pour le recrutement et la formation des enquêteurs. D'une façon générale, les enquêteurs ne sont pas recrutés et formés exclusivement pour les enquêtes HBS et LFS, mais sont aussi utilisés pour les autres enquêtes sur les ménages effectuées dans le pays. Dans toutes les enquêtes HBS, la collecte des données comporte un ensemble : *a)* de journaux tenus par les ménages ou les individus, généralement de façon quotidienne; et *b)* d'une ou de plusieurs entrevues.

29. Pour les enquêtes LFS, l'entrevue personnelle face à face est le principal mode de collecte des données. La « personne choisie comme référence » donne des informations sur le ménage, et chaque membre remplit un questionnaire individuel. Les entrevues d'autres personnes sont rares, mais la plupart des pays les considèrent comme une source de données valable. Lorsque le répondant ne peut être contacté personnellement, la plupart des pays l'autorisent à remplir seul le questionnaire. Cette formule est jugée préférable à une entrevue avec une autre personne. Étant donné le teneur du questionnaire, les entrevues téléphoniques ne sont pas largement utilisées, quoique l'on note certaines premières tentatives d'utilisation de la formule de l'en-

trevue téléphonique assistée par ordinateur [*computer-assisted telephone interviewing (CATI)*] (Estonie). La plupart des pays utilisent la formule classique « du crayon et du papier » pour les entrevues.

Le cadre d'échantillonnage pour les enquêtes HBS et LFS

30. Les recensements de la population et des logements servent de base à l'établissement du cadre d'échantillonnage des enquêtes sur les ménages de plusieurs pays (Bulgarie, Hongrie, Pologne et Roumanie). Les données des recensements sont utilisées pour créer les unités primaires d'échantillonnage (UPE) sur la base des unités d'énumération du recensement (UER), généralement adaptées aux exigences particulières de l'enquête. Dans la plupart des cas, les logements servent d'unités secondaires d'échantillonnage (USE). En général, les logements des UPE sélectionnées font l'objet d'une actualisation annuelle, qui comporte une estimation de l'accroissement du nombre de logements dû à l'achèvement de nouvelles constructions et une estimation de la diminution de ce nombre dû à la démolition d'immeubles et à des rectifications des limites des districts par suite de modifications du découpage administratif du pays [Główny Urząd Statystyczny (GUS), 1999; Kordos, 1982, 1996; Lednicki, 1982; Martini, Ivanova et Novosyolova, 1996; Mihalyffy, 1994],

31. Certains pays de l'ex-Union soviétique, par exemple le Bélarus, l'Estonie, la Lettonie et la Lituanie, utilisent les registres de population (RP) et les adresses indiquées sur ces registres ainsi que toute autre documentation administrative disponible comme cadres d'échantillonnage. (Lapins et Vaskis, 1996; Martini, Ivanova et Novosyolova, 1996; Šniukstiene, Vanagaite et Binkauskienė, 1996; Traat, Kukk et Sõstra, 2000).

32. Dans la Fédération de Russie, le microrecensement de 1994 a servi efficacement de cadre d'échantillonnage pour les enquêtes HBS et LFS (Goskomstat, 2000).

33. En général, la population cible couverte comprend tous les ménages du territoire national de chaque pays, à quelques rares exceptions mineures près. Parfois, certains petits groupes de population ne sont pas couverts, principalement par suite des limites de la couverture du cadre d'échantillonnage disponible.

34. On envisage à l'avenir d'utiliser les résultats des recensements de la population de 2000 comme cadres d'échantillonnage pour les enquêtes HBS et LFS et les autres enquêtes sur les ménages (Éltető et Mihalyffy, 2002; Kordos, Lednicki et Zyra, 2002; Lapins *et al.*, 2002; Kurvits, Sõstra et Traat, 2002).

Taille de l'échantillon et allocation

35. Pour 2000, la taille des échantillons pour les enquêtes HBS a varié de 1 028 ménages en Slovénie et 1 300 ménages en Slovaquie à 36 163 en Pologne et à 48 675 en Fédération de Russie. Le tableau XXV.2 indique les tailles des échantillons pour les enquêtes HBS et LFS de 2000 dans les pays en transition.

36. D'une façon générale, les grands pays, en raison de leur nécessité accrue de résultats fractionnés et de leur plus grande capacité, ont nécessité des échantillons de taille plus grande, mais bien entendu, pas en proportion de leur propre taille. Dans certains pays, l'échantillon a été distribué de façon proportionnelle à travers les régions géographiques, de manière à maximiser la précision des estimations à l'échelon national. Cependant, trois pays — la Hongrie, la Pologne et la Fédération de Russie — ont opté pour des allocations non proportionnelles, allouant aux régions plus petites des échantillons de proportion plus grande, de manière à assurer une taille minimum aux échantillons de chaque région du pays.

37. En 2000,
- a) *HBS* : la Fédération de Russie avait le plus gros échantillon (48 675 ménages, suivie de la Pologne (36 163), de l'Ukraine (12 534) et de la Hongrie (11 862). Les pays aux échantillons les plus petits étaient la Slovaquie (1 300) et la Slovénie (1 028);
 - b) *LFS* : la Fédération de Russie avait le plus gros échantillon (123 041), suivie de l'Ukraine (38 695), de la Hongrie (36 500 par trimestre), de la Pologne (24 400 par trimestre), de la République tchèque (31 800), de la Bulgarie (24 000), de la Roumanie (17 600) et de la Slovaquie (10 250). Tous les autres pays ont utilisé des échantillons de moins de 10 000 ménages.

Conception et sélection de l'échantillon

38. Au cours des 10 dernières années, les pays en transition ont utilisé des échantillons de conception différente pour leurs enquêtes *HBS* et *LFS*. Ils ont appliqué des critères différents pour la stratification des UPE avant la sélection. Les critères les plus courants étaient la région géographique et ou l'environnement urbain ou rural. Certains pays (notamment la Hongrie, la Pologne, la Fédération de Russie et l'Ukraine) ont utilisé la stratification en fonction de la population de la localité.

39. La plupart des enquêtes étaient fondées sur un échantillonnage à deux phases : la sélection des unités primaires d'échantillonnage (UPE) lors de la première phase, suivie, lors de la deuxième phase, de la sélection d'un petit nombre de logements ou de ménages au sein de chaque UPE sélectionnée. Normalement, les probabilités de sélection lors de ces deux phases étaient équilibrées de manière à obtenir un échantillon « autopondéré » à l'intérieur des domaines, c'est-à-dire que les UPE étaient sélectionnées selon une probabilité proportionnelle à leur taille, généralement le nombre de logements; et à l'intérieur des UPE sélectionnées, on choisissait le même nombre d'unités secondaires d'échantillonnage (USE). Dans les grandes villes de Lettonie et de Lituanie, des échantillons directs (une seule phase) de logements, de ménages ou de personnes ont été choisis. En revanche, en Hongrie, dans les petites localités, l'échantillon a été sélectionné en trois phases : les zones étendues durant la première phase, des grappes plus petites durant la deuxième et les adresses ou les ménages durant la dernière phase.

Renouvellement de l'échantillon

40. On peut alléger la charge d'avoir à répondre en procédant à des renouvellements périodiques des ménages de l'échantillon. Toutefois, ce renouvellement alourdit le coût de l'enquête car il nécessite un surcroît d'effort pour la tenue de l'échantillon, éventuellement un surcroît de formation d'enquêteurs, des coûts additionnels liés à la collecte d'informations de base et aux difficultés d'avoir à apprendre aux nouvelles unités à fournir les données demandées. Le renouvellement partiel des unités sélectionnées à un rythme fixe est choisi comme solution de compromis entre le renouvellement total, c'est-à-dire le remplacement de toutes les unités, qui est très coûteux et donne de mauvaises estimations du changement, et l'absence totale de renouvellement (autrement dit, une enquête par panel) qui aboutit à une distribution inacceptable de la charge de la réponse. Les formules de renouvellement maintiennent une unité dans l'échantillon pendant une période donnée, après quoi cette unité ne peut plus être sélectionnée pour la même enquête pendant un délai minimum.

41. La plupart des pays en transition appliquent une certaine formule de renouvellement à la fois aux enquêtes *HBS* et *LFS*. Par exemple, l'Estonie, la Pologne et la Roumanie ont appliqué la formule 2-(2)-2, c'est-à-dire deux trimestres dans l'échan-

tillon, deux trimestres hors de l'échantillon et deux autres trimestres dans l'échantillon, après quoi, les unités sont exclues.

Pondération des résultats

42. Les taux de non-réponse à l'enquête HBS sont généralement élevés, et ils varient considérablement selon la structure socio-économique des ménages de l'échantillon. Pour limiter l'impact de ce phénomène, on affecte des pondérations aux résultats de l'échantillon. On peut réduire sensiblement à la fois l'erreur d'échantillonnage et l'erreur autre que l'erreur d'échantillonnage si l'on dispose d'importantes informations auxiliaires et si on les utilise pour une nouvelle pondération selon une méthode d'étalonnage. La Hongrie est le seul pays où l'étalonnage a été utilisé pour les deux types d'enquêtes (Éltető et Mihalyffy, 2002; voir également Deville et Särndal, 1992).

43. Il peut être utile pour la conception et la sélection de l'échantillon de disposer d'informations sur les caractéristiques de base des unités du cadre. Plus encore, ces informations peuvent servir à calculer les pondérations, qui sont alors appliquées pour réduire l'effet de la non-réponse. Pour ce faire, il faut que l'on dispose à la fois des informations requises sur les caractéristiques des unités qui répondent comme de celles qui ne répondent pas.

44. Premièrement, chaque ménage de l'échantillon est pondéré selon l'inverse de la probabilité de sa sélection. La pondération pour les cas de non-réponse implique la division de l'échantillon en catégories de pondération appropriées et, à l'intérieur de chaque catégorie, les répondants se voient affecter une pondération qui tient compte des cas de non-réponse de cette catégorie. En outre, pour l'enquête HBS, on applique les pondérations appropriées de l'enquête LFS (pour la taille des ménages et la relation entre milieu urbain et rural) [Pologne].

45. Dans les États baltes, des procédures spéciales sont utilisées pour produire un échantillon autopondéré de ménages à partir d'un registre de la population (Lapins et Vaskis, 1996; Šniukstiene, Vanagaite et Binkauskienė, 1996; Traat, Kukk et Sõstra, 2000).

46. Les données de l'enquête LFS sont utilisées en même temps pour les analyses aux niveaux des ménages et des personnes. Il est donc nécessaire d'utiliser une procédure de pondération qui assure une parfaite homogénéité dans l'analyse portant sur ces deux types d'unités. Toute pondération de l'échantillon initial est appliquée au niveau des ménages, ce qui veut dire que la procédure appliquée garantit la même pondération à tous les membres d'un même ménage.

47. Les pondérations sont établies à tour de rôle. À chaque étape après la première, les pondérations sont calculées à partir de valeurs déjà pondérées de l'échantillon selon les résultats des étapes précédentes. La pondération finale d'une unité est le produit des coefficients de pondération déterminés à chaque étape. Les pondérations calculées à chaque étape sont normalisées, c'est-à-dire qu'elles sont calibrées de telle manière que la valeur moyenne par unité de l'échantillon est égale à 1,0 et que la somme des pondérations est égale à la taille initiale de l'échantillon.

48. Bien qu'un ensemble commun de procédures soit utilisé dans toutes les enquêtes, les variables spécifiques intervenant à chaque étape et les sources des données utilisées varient d'une enquête à l'autre. Néanmoins, certaines variables tendent à être importantes dans presque tous les cas, notamment la localisation géographique du ménage, sa taille et sa composition, et la répartition de la population par âge, sexe et autres caractéristiques essentielles (Verma, 1995),

Tableau XXV.2
**Taille de l'échantillon et conception et méthodes d'estimation de l'échantillon
dans les enquêtes HBS et LFS de 2000 de certains pays en transition**

Pays	Taille de l'échantillon		Conception de l'échantillon		Méthode d'estimation	
	HBS	LFS	HBS	LFS	HBS	LFS
Bélarus	6 000	–	2-phases	–	Pondérée	–
Bulgarie	6 000	24 000	2-phases PPS	2-phases PPS	Directe	Pondérée
Croatie	2 865	12 843	2-phases PPS	2-phases PPS	Pondérée	Direct
République tchèque	3 250	31 800	Quota	2-phases PPS	Dernier microc.	Pondérée
Estonie	9 840	9 127	PR PPS	PR Eq. Pr.	Pondérée	Pondérée
Hongrie	11 862 ^a	36 500 ^b	3-phases PPS	3-phases ^c PPS	Calibr.	Calibr.
Lettonie	3 847	7 940	2-phases PPS	2-phases PPS	Pondérée	Pondérée
Lituanie	10 680	6 000	PPS personne	PPS personne	Pondérée	Pondérée
Pologne	36 163 ^d	24 400 ^e	2-phases PPS	2-phases PPS	Pondérations LFS	Pondérations demogr.
Roumanie	17 827	17 600	2-phases PPS	2-phases PPS	Pondérée	Pondérée
Fédération de Russie	48 675	123 041	2-phases PPS	2-phases PPS	Microc. pondéré	microc. Pondéré
Slovaquie	1 300	10 250	quota	2-stage PPS	–	Pondérée
Slovénie	1 028	7 000 ^f	2-phases PPS	1-stage personne	Pondérée	Pondérée
Ukraine	12 534	38 695	2-phases PPS	2-phases PPS	Pondérée	Pondérée

Source : Données tirées des questionnaires fournis par certains pays.

Note : « Dernier microc. » signifie le microrecensement de 1995; « microc. pondéré » = microcensus pondéré; « PR » = Registre de population; « Pondérations LFS » = pondérations utilisées dans l'enquête LFS; « Pondérations demogr. » = pondérations utilisées dans les projections démographiques-données de contrôle post-stratification; « calibr. » = méthode de calibration; « Eq. Pr. » = probabilité égale; tiret (–) = données sans objet.

^a Le nombre de ménages ayant coopéré à l'enquête a été de 10 191. Pour obtenir ce résultat, l'enquêteur a dû utiliser 17 243 adresses.

^b Sélection trimestrielle.

^c Sauf pour les villes autoreprésentées, où la sélection s'est faite en deux phases.

^d Cette taille d'échantillon n'a été obtenue qu'en 2000. L'année précédente, l'échantillon était d'environ 32 000 ménages.

^e Nombre trimestriel de logements choisis. Chaque trimestre, le même nombre de logement est sélectionné.

^f Chiffre trimestriel.

Estimation des erreurs types

49. La plupart des pays utilisent des échantillons de conception complexe pour les enquêtes HBS et LFS et sont donc tenus d'incorporer ces éléments complexes dans le calcul de la variance d'échantillonnage (Wolter, 1985). Les expressions analytiques de la variance ne permettent pas d'estimer l'erreur d'échantillonnage d'estimations compliquées, de sorte que l'on utilise des méthodes par approximation. Certains pays ont utilisé la méthode du groupe aléatoire (c'est notamment le cas de la Pologne pour l'enquête HBS jusqu'en 2000 et pour l'enquête LFS jusqu'en 1999), la méthode « jack-knife » (Hongrie), la méthode de la série de Taylor (Pologne pour l'enquête LFS depuis

le quatrième trimestre de 1999), la méthode du demi-échantillon équilibré (Pologne pour l'enquête HBS depuis 2001) et une méthode analytique adaptée au pays (Fédération de Russie). Certains pays (l'Estonie, la Lettonie et la Slovénie) s'appuient sur le logiciel pour l'analyse statistique de données corrélées (SUDAAN), le logiciel bien connu utilisé pour calculer les erreurs types sur les échantillons de conception complexe.

Taux de non-réponse dans les enquêtes HBS et LFS

50. Si nous prenons les taux moyens de non-réponse enregistrés ces quatre dernières années dans les enquêtes HBS de certains pays en transition, nous pouvons identifier les trois groupes suivants tels qu'ils ressortent des données du tableau XXV.3.

- a) Le groupe à taux élevé de non-réponse (supérieur à 40 %) : Estonie (43,6 %), Pologne (43,4 %), Bulgarie (41,7 %) et Hongrie (40,0 %);
- b) Le groupe à taux intermédiaire de non-réponse (entre 20 % et 30 %) : Fédération de Russie (25,6 %), Ukraine (25,0 %), Lettonie (24,5 %), et Lituanie (22,2 %);
- c) Le groupe à faible taux de non-réponse (moins de 20 %) : Croatie (19,0 %), Slovénie (18,5 %) et Roumanie (11,0 %).

Tableau XXV.3

Taux de non-réponse à l'enquête HBS dans certains pays en transition, 1992-2000

Pays	Taux de non-réponse de certaines années								
	1992	1993	1994	1995	1996	1997	1998	1999	2000
	<i>en pourcentage</i>								
Bulgarie	..	33,0	34,2	35,6	37,9	49,0	41,1	39,7	37,0
Croatie	19,0	21,0	17,0
République tchèque	Non indiqué								
Estonie	44,4	50,2	44,9	46,6	47,5	35,2
Hongrie	..	36,7	40,4	32,6	43,3	40,6	40,9	39,6	39,0
Lettonie	26,1	24,1	21,9	23,1	28,7
Lituanie	24,0	20,3	22,7	22,8	22,8
Pologne	23,2	27,6	25,3	25,1	31,4	34,3	40,7	49,4	49,2
Roumanie	8,0	10,2	9,6	10,4	11,6	13,4
Fédération de Russie	10,4	10,5	5,9	11,5	31,4	47,5	25,0	13,9	16,0
Slovaquie	Non indiqué								
Slovénie	..	24,6	22,1	28,0	34,6	19,5	18,4	17,6	18,6
Ukraine	24,2	25,7

Source : Questionnaires spéciaux par pays.

Note : Deux points (..) indiquent l'absence de données.

51. Comme le montrent les tableaux XXV.3 et XXV.4, les taux de non-réponse dans les enquêtes HBS sont beaucoup plus élevés que ceux des enquêtes LFS dans tous les pays. De plus, dans certains pays, on note un net accroissement des taux de non-réponse dans le temps pour ces deux types d'enquêtes. Pour les enquêtes LFS, on peut noter un certain accroissement du taux de non-réponse dans les pays suivants :

- a) Pologne (4,5 % en 1992 contre 22,1 % en 2000);
- b) Bulgarie (10,1 % en 1993 contre 17,2 % en 2000);
- c) République tchèque (16 % en 1993 contre 24 % en 2000);

- d) Croatie (6,3 % en 1996 contre 15,7 % en 2000);
- e) Roumanie (2,6 % en 1994 contre 8,9 % en 2000);
- f) Slovénie (9,0 % en 1992 contre 12,0 % en 2000).

52. Les données du tableau XXV.4 indiquent que les taux de non-réponse dans les enquêtes LFS diffèrent considérablement d'un pays à un autre et peuvent se diviser en trois groupes fondés sur le niveau de non-réponse :

- a) Le groupe à taux élevé de non-réponse (supérieur à 15 %) : Ukraine (28,8 %), République tchèque (21,5 %), Bulgarie (16,1 %), Croatie (15,7 %) et Pologne (15,4 %);
- b) Le groupe à taux intermédiaire de non-réponse (entre 10 et 15 %) : Estonie (12,5 %), Slovénie (12,2 %), Hongrie (11,2%) et Lettonie (10,4 %);
- c) Le groupe à faible taux de non-réponse (moins de 10 %) : Lituanie (9,1 %), Roumanie (7,7 %), Slovaquie (5,6 %) et Fédération de Russie (5,4 %).

Tableau XXV.4

Taux de non-réponse à l'enquête LFS dans certains pays en transition, 1992-2000

Pays	Taux de non-réponse de certaines années								
	1992	1993	1994	1995	1996	1997	1998	1999	2000
	<i>en pourcentage</i>								
Bulgarie	..	10,1	8,8	8,5	11,0	14,4	16,7	16,2	17,2
Croatie	6,3	14,0	18,1	15,0	15,7
République tchèque	..	16	16	18	20	19	21	22	24
Estonie	7,4	..	13,5	13,4	13,2	9,9
Hongrie	..	10,3	8,1	11,4	13,6	14,3	12,2	8,9	9,2
Lettonie	13,7	13,3	12,4	9,8	9,4	10,1
Lituanie	9,6	9,0	8,7	8,9
Pologne	4,5	5,3	8,9	9,9	10,0	9,6	11,6	18,2	22,1
Roumanie	2,6	2,3	6,4	6,7	7,4	7,9	8,9
Fédération de Russie	4,6	6,8	5,9	4,5	5,5	5,8	5,8	5,3	4,5
Slovaquie	6,2	5,9	5,1	5,0	5,6	5,9	5,7
Slovénie	9,0	7,9	9,8	9,7	10,0	12,5	12,4	11,7	12,0
Ukraine	29,2	28,3

Source : Questionnaires spéciaux par pays.

Note : Deux points (..) indiquent l'absence de données.

Coûts des enquêtes sur les ménages

53. Dans toute enquête sur les ménages, il faut répondre à deux questions importantes, à savoir :

- a) Quel est le coût total de l'enquête ?
- b) Quel est le degré de précision des principales estimations ?

54. Il n'est pas facile d'estimer les coûts des enquêtes sur les ménages dans les pays en transition. Certains pays ne donnent que le coût total direct de la collecte de données, y compris les entrevues, les déplacements, les matériels et les services liés à cette collecte, mais non compris certaines autres composantes telles que la préparation de l'enquête, les moyens d'imputation méthodologique, le traitement des données et la rédaction et la publication des rapports.

55. Malgré l'importance cruciale de l'établissement d'un budget, l'estimation des coûts est l'un des aspects les moins développés de la planification des enquêtes. L'un des problèmes que comporte cette estimation a trait au caractère souvent laborieux de la tenue d'une comptabilité détaillée. L'autre est qu'il est difficile de ventiler les coûts d'efforts conjoints, surtout les dépenses administratives et autres coûts indirects. Néanmoins, la mise en place d'un système détaillé de notification des dépenses pourrait contribuer pour beaucoup aux progrès de la planification à venir et à l'aptitude à attirer l'appui nécessaire aux programmes de collecte de données (Nations Unies, 1984).

56. En Pologne (Kordos, Lednicki et Zyra, 2002), le coût direct de l'enquête HBS de 2000 a été de 4 567 000 € dont 3 571 394 € (78,2 %) pour les entrevues, 146 144 € (3,2 %) pour les frais de déplacement et 429 298 € (9,4 %) pour les incitations. Comme en 2000, la taille de l'échantillon des ménages interrogés était de 36 163, ce qui représente un coût moyen par ménage de 126,3 €.

57. Des calculs similaires ont été effectués pour l'enquête LFS de 2000. Les coûts directs de cette enquête ont été de 1 094 200 €, dont 878 642,6 € (80,3 %) pour les entrevues et 45,956,4 € (4,2 %) pour les frais de déplacement. L'enquête LFS n'avait pas prévu de coûts d'incitation. Étant donné qu'en 2000, près de 80 000 ménages ont été interrogés, chaque entrevue a coûté en moyenne 13,7 €. On remarquera que le coût de l'enquête HBS a été d'environ 10 fois supérieur à celui de l'enquête LFS, ce qui s'explique principalement par le fait que l'enquête HBS demande beaucoup de temps et nécessite plusieurs entrevues avec les mêmes répondants ainsi que la tenue d'un journal et de pièces justificatives, alors que l'enquête LFS ne nécessite qu'une entrevue.

58. La Hongrie a fourni des données intéressantes sur les coûts des enquêtes HBS et LFS de 2000 (Éltető et Mihalyffy, 2002). Les tableaux XXV.5 et XXV.6 présentent des évaluations détaillées de la structure des coûts des enquêtes HBS et LFS effectuées en 2000 dans ce pays. Les dépenses relatives à l'enquête LFS (432 000 €) ont dépassé celles de l'enquête HBS (326 000 €). Toutefois, si l'on considère que pour l'enquête LFS, le même ménage a dû être interrogé quatre fois dans l'année et qu'aucune incitation n'a été donnée aux ménages qui acceptaient de coopérer, les dépenses par ménage ont été nettement plus faibles que pour l'enquête HBS (27,5 € par ménage pour l'enquête HBS, contre 8,4 € par ménage pour l'enquête LFS). Les tableaux XXV.5 et XXV.6 donnent la structure des coûts des enquêtes HBS et LFS à la fois en valeur absolue (€) et en pourcentages.

tableau XXV.5

Structure des coûts de l'enquête HBS de 2000 en Hongrie

Élément de coût	Coût en	
	Euros	Pourcentage
Coût mensuel des journaux	148 650	45,6
Questionnaires à remplir en fin d'année	35 865	11,0
Visite aux ménages non répondants	4 345	1,3
Incitations données aux ménages qui coopèrent	75 855	23,3
Prime aux enquêteurs	18 585	5,7
Coûts du matériel	42 700	13,1
Total	326 000	100,0

Source : Éltető et Mihalyffy (2002).

Tableau XXV.6
Structure des coûts de l'enquête LFS de 2000 en Hongrie

Élément de coût	Coût en	
	Euros	Pourcentage
Visites aux ménages	22 032	5,1
Questionnaires à remplir par les ménages	65 232	15,1
Questionnaires sur les activités	212 110	49,1
Questionnaires supplémentaires	42 336	9,8
Prime aux enquêtes	33 695	7,8
Coûts du matériel	56 595	13,1
Total	432 000	100,0

Source : Éltető et Mihályffy (2002).

Effets de conception

59. Comme le montre la description des enquêtes sur les ménages effectuées dans certains pays en transition, presque toutes ces enquêtes sont fondées sur une structure à phases multiples. Autrement dit, il faut calculer les effets de conception pour l'analyse statistique des données de ces enquêtes (Kish et Frankel, 1974).

60. Nous présentons un exemple tirée de la Fédération de Russie (Goskomstat, 2000, pp. 219-220), où la taille de l'échantillon de l'enquête LFS trimestrielle a été déterminée séparément pour chaque région du pays. Cette taille a été déterminée pour divers niveaux de taux de chômage réel. Le niveau de précision désiré de l'estimation a été fixé à 1,5 %; à 5 % et à 8 % respectivement pour l'ensemble du pays, pour les régions grandes et moyennes et pour les petites régions. L'effet de conception a été calculé selon la formule de l'équation (7) du chapitre VI, et sur la base des données de l'enquête par sondage sur l'emploi et le chômage de 1998. Les effets de conception calculés se situaient dans une fourchette de 1,52 à 2,14. Ils ont été calculés pour plusieurs caractéristiques des enquêtes HBS et LFS.

7. Conclusions

61. Dans ce chapitre, nous avons présenté divers aspects de la conception et de l'exécution d'enquêtes par sondage sur les ménages, nous concentrant sur les enquêtes les plus importantes : les enquêtes HBS et LFS. Cet examen général des enquêtes sur les ménages nous permet de tirer certaines conclusions. Les enquêtes par sondage sur les ménages dans les pays en transition ont été restructurées et harmonisées conformément aux nouvelles exigences de l'économie de marché et aux recommandations d'Eurostat (1995; 1996; 1997; 1998a), avec certaines différences entre pays liées à l'expérience passée et aux possibilités actuelles. Bien que les progrès aient été évidents en ce qui concerne l'évolution de ces enquêtes, un certain nombre de problèmes nécessitaient un examen plus poussé au niveau du bureau de statistique, notamment pour le calcul et la présentation des erreurs types, l'évaluation des éléments de coût et le calcul et la publication des effets de conception et leur utilisation dans l'analyse statistique. Par ailleurs, certains pays se heurtent à des problèmes particuliers, tels que de faibles taux de réponse et un échantillon de taille inadéquate pour certains domaines. Ces problèmes sont très importants et graves pour la comparabilité des résultats entre pays. Il incombe à Eurostat de les résoudre, dans la mesure où ils nuisent à l'intégra-

tion et à l'harmonisation des enquêtes par sondage sur les ménages effectuées dans divers pays. Les pays en transition ont leurs propres plans pour la poursuite du développement de leurs enquêtes sur les ménages. L'un de ces plans implique l'utilisation des résultats des recensements de la population et des logements de 2000. Les données de ces recensements offrent la possibilité d'améliorer les cadres d'échantillonnage, la conception des échantillons et les méthodes d'estimation, surtout pour les domaines de petite taille.

62. Les diverses études de cas tirées de certains pays en transition qui sont présentées ci-dessous offrent un tableau plus détaillé des problèmes de conception, d'exécution et d'analyse des enquêtes sur les ménages dans ces pays. Ces études de cas sont suivies d'une liste détaillée de références, qui peuvent être utiles pour étudier divers aspects des enquêtes sur les ménages dans les pays en transition.

B. Les enquêtes sur les ménages dans les pays en transition : études de cas

63. Les études de cas présentées dans cette section ont été préparées par les auteurs et portent sur les pays suivants : Estonie, Hongrie, Lettonie, Lituanie, Pologne et Slovaquie. Des articles plus détaillés sur huit pays en transition ont été publiés dans *Statistics in Transition*, vol. 5, n° 4 (juin 2002). Les informations sur les principales caractéristiques des enquêtes HBS et LFS et d'autres enquêtes sur les ménages viennent compléter les informations données dans la section A.

1. L'enquête par sondage sur les ménages en Estonie²

Introduction

64. En 1995, le Bureau de statistique de l'Estonie a exécuté deux grandes enquêtes sur les ménages : l'Enquête sur la main-d'œuvre estonienne (LFS) et l'Enquête sur le budget des ménages estoniens (HBS).

65. L'enquête HBS est continue : ses résultats sont publiés trimestriellement et annuellement. En 1999, elle a été restructurée dans le cadre d'un projet de la Banque mondiale; ses journaux ont été modifiés et les procédures d'échantillonnage et de pondération ont été plus étroitement alignées sur les données les plus récentes. La description de cette enquête est donnée dans Traat, Kukk et Sõstra (2000) et de façon plus détaillée dans Traat (1999).

66. L'enquête LFS était une enquête ponctuelle effectuée en 1995. La version suivante, de 1997, a suivi une méthodologie différente. Après cela, elle a été trimestrielle jusqu'en 2000, où elle est devenue continue. Cette enquête est décrite dans Kurvits, Sõstra et Traat (2002) et dans une publication du Bureau de statistique de l'Estonie (1999).

67. Par ailleurs, le Bureau de statistique de l'Estonie a effectué de nombreuses autres enquêtes sur les ménages ou sur la population. Ces enquêtes correspondent à une série d'études similaires effectuées dans d'autres pays européens, et les informations obtenues ont permis de faire des comparaisons nationales et internationales. Il s'agit de l'Enquête sur l'éducation des adultes (1997), l'Enquête sur l'emploi du temps (1999-2000), l'Enquête sur les conditions de vie (1994 et 1999), et l'Enquête de 2000 sur le comportement de la population adulte estonienne en matière de santé (Kurvits, Sõstra et Traat, 2002).

68. Les enquêtes sur le budget des ménages et sur la main-d'œuvre de l'Estonie sont décrites brièvement ci-dessous.

² Préparée par Imbi Traat, Institut de statistiques mathématiques, Université de Tartu (courriel : imbi@ut.ee).

Les données

69. L'enquête HBS est fondée sur la tenue d'un journal. Chaque ménage sélectionné fournit des informations sur sa consommation et ses dépenses alimentaires d'une semaine, ainsi que sur ses autres dépenses et ses revenus d'un mois. Il est également procédé à une entrevue préalable, qui a pour but de recueillir des informations sur la composition du ménage et sur d'autres questions d'ordre général, et à une brève entrevue ultérieure sur les éventuels changements dans la composition du ménage.

70. Le programme de collecte de données de l'enquête LFS a été plus large que celui d'une enquête ordinaire sur la main-d'œuvre, surtout au cours de la période 1995-1999 où des informations rétrospectives ont été recueillies. La situation du répondant du point de vue de l'emploi (employé, chômeur ou inactif) a été enregistrée pour toute la période écoulée depuis l'enquête précédente. Les dates de commencement et de fin et toutes autres données pertinentes ont été consignées pour chaque situation nouvelle. Le module type des enquêtes sur la main-d'œuvre est appliqué à la semaine de référence et pose des questions aux personnes employées sur leur profession, leur durée de travail habituelle et effective, l'activité économique de l'entreprise/organisation, etc. Les personnes au chômage sont interrogées sur ce qu'elles ont fait pour trouver un emploi, leur persévérance dans la recherche d'un emploi et les caractéristiques de l'emploi qu'elles recherchent, etc.

Collecte des données

71. Le Département des enquêteurs (créé en 1994) du Bureau de statistique de l'Estonie est chargé de la collecte des données d'une variété d'enquêtes. Les 15 coordonnateurs de comté organisent le travail de 130 enquêteurs, répartis dans tout le pays. Dans les zones rurales, un enquêteur effectue en moyenne 10 à 15 entrevues par enquête en un mois, et dans les zones urbaines, environ 15 à 20, mais en réalité, leur charge de travail varie selon la taille des échantillons régionaux. Les entrevues sont le second emploi d'environ la moitié des enquêteurs. Ceux-ci sont rémunérés pour chaque entrevue menée à bien ainsi que pour les tentatives de prise de contact avec des non répondants.

72. L'entrée et la codification des données se font au Bureau de statistique selon le système appelé « Blaise ». Le programme d'entrée des données comporte la première vérification logique de celles-ci. Le traitement des données et les vérifications plus complexes sont effectués au moyen du logiciel « Système d'analyse statistique (SAS) » dans le cas de l'enquête LFS et du logiciel FoxPro dans celui de l'enquête HBS.

73. Le coût total des enquêtes HBS et LFS de 2000 a été respectivement de 153,000 € et de 128,000 €. Le salaire des enquêteurs, les transports et les communications représentaient environ 70 % de ce total, et l'entrée des données, environ 15 %.

Conception de l'échantillon

74. La population cible de l'enquête HBS regroupe tous les ménages estoniens, à l'exclusion des personnes vivant en institution. La population cible de l'enquête LFS est composée des habitants de l'Estonie âgés de 15 à 74 ans.

75. Le cadre d'échantillonnage des deux enquêtes est le Registre de la population. Les unités d'échantillonnage sont les personnes, et elles sont sélectionnées systématiquement à partir des dossiers de ce Registre. Des strates (trois pour l'enquête HBS

et quatre pour l'enquête LFS) présentant des taux d'échantillonnage différents sont utilisées pour permettre une meilleure couverture.

76. Les deux enquêtes ont utilisé des informations auxiliaires tirées du cadre. La fréquence d'une adresse dans le cadre détermine la probabilité d'inclusion de cette adresse. L'échantillon est divisé en deux parties, qui obéissent à des règles différentes : l'échantillon d'adresses (les dossiers contenant les adresses complètes) et l'échantillon de personnes (les dossiers dont les adresses sont incomplètes ou inconnues). Les adresses incomplètes ou inconnues correspondent aux régions rurales où l'adresse est simplement le nom du village, sans autre indication.

77. Dans l'échantillon d'adresses, tous les ménages vivant aux adresses sélectionnées sont incluses dans l'échantillon. Dans l'échantillon de personnes, seuls figurent dans l'échantillon les ménages dont un membre a été sélectionné. Le ménage est identifié dans un comté. Environ 15 % des ménages sont sélectionnés à travers des personnes sélectionnées.

78. La conception de l'échantillon utilise une probabilité proportionnelle à la taille [*probability-proportional-to-size* (PPS)], où la taille est soit la fréquence de l'adresse dans le cadre soit la taille du ménage (indiquée par celui-ci). Pour l'enquête HBS, cet échantillon est l'échantillon final. Sa probabilité PPS est utilisée pour déterminer les estimateurs. Pour l'enquête LFS, il s'agit d'un échantillon de première phase de ménages/d'adresses dont le nombre de membres en âge de travailler, s'il ne peut être tiré du Registre, est déterminé avec l'aide des autorités locales. Le but de l'échantillonnage de deuxième phase est de produire un échantillon de ménages (et de leurs membres) de probabilité égale. Tous les ménages (adresses) qui comptent un membre en âge de travailler et, par échantillonnage systématique, la moitié des ménages qui en comptent deux, un tiers des ménages qui en comptent trois, etc. sont inclus dans l'échantillon final.

79. Actuellement, l'enquête HBS sélectionne chaque mois 820 ménages.

80. Depuis 2000, les ménages inclus dans l'enquête LFS sont renouvelés selon un plan de renouvellement de 2-2-2. Les ménages sont interrogés à quatre reprises : pendant deux trimestres consécutifs et, après un hiatus de deux trimestres, pendant les deux trimestres correspondants de l'année suivante. D'après ce plan de renouvellement, chaque trimestre, 25 % des ménages participent pour la première fois à l'enquête et 50 % sont des ménages qui ont déjà été interrogés le trimestre précédent. Ainsi, il se produit un chevauchement de 50 % entre trimestres voisins et entre les mêmes trimestres d'années successives.

Non-réponse

81. Dans les enquêtes utilisant un journal, l'accroissement de la charge de la réponse tend à engendrer des taux de non-réponse plus élevés (voir section A du présent chapitre).

82. En général, les « refus » représentent environ 50 % du total de cas de non-réponse et les « absents », environ 25 %.

83. Le taux de non-réponse à l'enquête LFS est toujours beaucoup plus faible que celui de l'enquête HBS. Par ailleurs, le taux de refus est en hausse en raison du temps limité imposé pour le travail de terrain du fait du passage à l'approche de l'enquête continue. De plus, le fait d'inclure quatre fois les ménages dans l'enquête LFS favorise un accroissement des cas de non-réponse.

Pondération

84. La pondération est utilisée à la fois pour les enquêtes HBS et LFS. Dans les enquêtes HBS, elles sont calculées par ménage, et dans les enquêtes LFS, par personne.

85. Dans l'enquête HBS, les six groupes de pondération sont déterminés en fonction des taux de réponse et des niveaux de revenus/dépenses. Les pondérations initiales sont multipliées par l'inverse des taux de réponse du groupe. Les pondérations des réponses sont ensuite calibrées en fonction de la répartition par sexe/âge (cinq catégories) sur la base des statistiques démographiques connues.

86. Dans l'enquête LFS, les pondérations sont formulées en une série d'étapes (Verma, 1995). La pondération initiale d'un répondant est la taille de la population cible (personnes âgées de 15 à 74 ans) divisée par le nombre de répondants calculé dans chacune des quatre strates. Ensuite, on forme six groupes régionaux de pondération de taille raisonnablement uniforme avec des taux de réponse R_j différents. À l'intérieur de chaque groupe, le coefficient de correction de la pondération d'un répondant individuel est $w_j^{(0)} = \bar{R} / R_j$, où \bar{R} est le taux de réponse global (moyen). Après cela, on utilise la méthode du ratissage avec cinq itérations pour calibrer les distributions de l'échantillon selon les catégories de population en utilisant le sexe, l'âge (cinq groupes d'âge) et le lieu d'habitation (15 comtés et la capitale).

Paramètres et estimateurs

87. La plupart des paramètres estimés dans les enquêtes HBS et LFS sont des totaux et des ratios. On utilise les estimateurs pondérés Horvitz-Thompson ou leurs ratios.

88. Les estimations de la variance sont utilisées à l'aide de SUDAAN. Comme ce logiciel ne peut utiliser la conception exacte des enquêtes LFS et HBS, on utilise la conception la plus proche possible – échantillonnage en grappes stratifié avec remplacement probabilité inégale, avec les ménages comme grappes. En raison de l'hypothèse de l'échantillonnage avec remplacement, les estimations tendent à surestimer légèrement les vraies variances.

Perspectives d'avenir

89. Le Recensement de la population de 2000 fournit une mine d'informations sur les ménages et les individus estoniens. Le système de pondération des enquêtes HBS et LFS sera revu à la lumière des données de ce recensement qui reflètent mieux la situation démographique de l'Estonie que les données utilisées auparavant.

90. On s'efforcera également d'améliorer d'autres phases des enquêtes. Par exemple, en 2002, une nouvelle méthode de collecte de données — entrevues téléphoniques assistées par ordinateur [*computer-assisted telephone interviewing* (CATI)] à l'aide de 10 ordinateurs portables — a été expérimentée pour l'enquête LFS. Il est prévu un essai d'entrevues face à face au premier contact et d'entrevues téléphoniques pour les trois entrevues suivantes.

2. Conception et exécution de l'Enquête sur le budget des ménages et de l'Enquête sur la main-d'œuvre en Hongrie³

Enquête sur le budget des ménages

91. L'enquête HBS obéit à une vieille tradition en Hongrie. Elle a commencé dans les années 1950, d'abord sur la base d'échantillons par la méthode des quotas.

³ Préparé par Ödön Éltető et László Mihályffy, Central Statistical Office, P.O. Box 51, H-1525 Budapest, Hongrie.

Puis elle a utilisé un échantillon probabiliste dans le cadre du Système unifié d'enquêtes sur les ménages, vers le milieu des années 1970. Le cadre d'échantillonnage a toujours été constitué des districts d'énumération (DE) du recensement, actualisés après chaque recensement décennal, le plus récent remontant à 2002. Entre 1976 et 2002, l'enquête HBS a été continue; entre 1983 et 1991, elle a été effectuée tous les deux ans; et depuis 1993, elle est à nouveau continue.

92. L'échantillon de l'enquête HBS est sélectionné en trois phases, sauf pour les villes qui se représentent elles-mêmes, c'est-à-dire les villes de 7 000 logements ou plus, où le processus de sélection ne comporte que deux phases. Dans le cas des localités qui ne se représentent pas elles-mêmes, les unités primaires d'échantillonnage (UPE) sont les localités, les unités secondaires sont les districts d'énumération (DE) et les ultimes unités d'échantillonnage sont les logements. Dans les villes qui se représentent elles-mêmes, les DE sont les UPE.

93. Les localités sont stratifiées selon leur taille, produisant huit strates, et aussi par comté. Généralement, l'échantillon n'est pas alloué proportionnellement aux strates. Le taux d'échantillonnage est plus faible dans les petites localités que dans les grandes villes, en particulier Budapest. La taille de l'échantillon annuel est distribuée de façon homogène sur les douze mois.

94. Il est demandé au ménage qui consent à participer à l'enquête d'indiquer son revenu et ses dépenses de façon journalière sur une période d'un mois. Pendant cette période, les enquêteurs collectent des données supplémentaires sur le ménage, notamment l'âge et la structure professionnelle du ménage, le type, la taille et l'équipement de son logement, les biens de consommation durable dont il dispose, etc. De plus, au début de l'année suivante, l'enquêteur rend de nouveau visite au ménage et interroge ses membres sur les grosses dépenses moins fréquentes qu'il a effectuées au cours de l'année et sur certains types de revenu annuel.

95. Le fait que, tous les deux ans, les enquêteurs rendent visite à chaque ménage de leur DE pour recueillir des informations démographiques et économiques, telles que la taille du ménage, l'âge, le niveau d'instruction et l'activité économique de son chef, est un important aspect de l'enquête HBS. Ces données sont utilisées principalement aux fins de substitution : du fait du taux de non-réponse assez élevé, l'utilisation de ménages de substitution (deux dans les grandes villes et un ailleurs) est autorisée. Le ménage de substitution est choisi dans la même strate que le ménage choisi à l'origine et dans le même DE attribué à l'origine à l'enquêteur.

96. Chaque année, le renouvellement d'un tiers des ménages est tel que la taille de l'échantillon à l'intérieur de chaque DE reste constante (six logements). Comme il peut arriver que les ménages de substitution ne répondent pas, le nombre effectif de ménages coopérant à l'enquête peut être plus grand ou plus petit que le nombre initial de six. Autrement dit, le rythme de renouvellement à l'intérieur d'un DE donné peut être supérieur ou inférieur à un tiers. Tout ménage qui a participé à l'enquête trois années de suite est définitivement remplacé.

97. En 2000, l'échantillon de l'enquête HBS a porté sur près de 1 980 DE pris dans 262 localités, et le nombre de ménages initialement sélectionnés a été de 11 862.

98. Comme les enquêteurs se heurtent souvent aussi à des refus ou à d'autres types de non-réponse aux adresses de substitution, le nombre définitif d'entrevues est inférieur à la taille prévue de l'échantillon. Ainsi, en 2000, au lieu de 11 862 ménages ($1\,977 \times 6$), seuls 10 191 ont participé à l'enquête, et pour obtenir ce résultat, les enquêteurs ont dû frapper à 17 243 adresses. Les taux de non-réponse ont augmenté après 1993, atteignant 43,3 % en 1996, puis ont diminué légèrement. En 2000, le taux de non-réponse a été de 39 % et les refus représentaient près de 27 % de ce total, le

reste étant constitué de logements vacants, de ménages absents, d'adresses inexactes et autres. Étant donné le problème que représentent les cas de non-réponse, il est particulièrement difficile d'atteindre la taille prévue de l'échantillon dans la capitale et certaines autres grandes villes. Bien que, jusqu'à la fin de 2002, les ménages participants aient reçu une incitation monétaire pour les informations fournies, le montant versé n'était pas suffisant pour inciter de nombreux ménages à coopérer à l'enquête, et il a donc été mis fin à ce programme d'incitations. En revanche, le régime de rémunération des enquêteurs a pris un tour favorable, les encourageant à redoubler d'efforts pour persuader les ménages de coopérer. Globalement, le taux de refus est tombé de 34,4 % en 1996 à 26,9 % en 2000.

99. La conception de l'échantillon de l'enquête HBS remplit les conditions nécessaires à l'utilisation de l'estimateur connu Horvitz-Thompson. Les totaux sont des sommes pondérées des observations, et les pondérations sont l'inverse des probabilités d'inclusion. Dans chacune des 98 strates de l'échantillon HBS, la pondération est unique et définie comme le ratio du nombre de logements non vacants de la strate au nombre d'entrevues menées à bien. En raison de la non-réponse unitaire et d'éventuelles déficiences de couverture, les pondérations ne permettent pas de calculer les données, et il faut donc utiliser des pondérations calibrées. Dans le processus de calibrage, les pondérations sont ajustées à l'aide des variables auxiliaires suivantes :

- Groupe âge-sexe (2 × 4 catégories);
- Activité économique (9 catégories);
- Niveau d'instruction (3 catégories);
- Type de ménage (3 catégories).

100. Dans le cas des données trimestrielles, le calibrage est effectué pour trois grands territoires : la capitale, les villes ayant des droits d'un comté et le reste du pays. Pour les données annuelles, le découpage territorial utilisé pour le calibrage est plus complexe. Les sept régions du pays — les régions du niveau II d'Eurostat de la Nomenclature des unités territoriales statistiques (NUTS) — sont également prises en compte.

101. Les pondérations calibrées de l'enquête HBS sont calculées selon la procédure généralisée de pondération du ratio de ratissage.

102. Les estimations des erreurs d'échantillonnage sur les revenus et les dépenses tirés des données de l'enquête HBS sont régulièrement calculées et publiées. Pour ces calculs, on utilise l'option « jackknife » stratifiée du logiciel VPLX mis au point par R. E. Fay. Pour l'avenir, on envisage d'utiliser la méthode « bootstrap », en particulier pour les estimations de quantiles.

103. L'enquête HBS est l'une des enquêtes les plus coûteuses du Bureau central de statistique [*Central Statistical Office (CSO)*]. En 2000, les dépenses directes liées à cette enquête; à l'exclusion des salaires du personnel du siège et des bureaux locaux du CSO, à savoir la rémunération des enquêteurs, les incitations des ménages à la coopération et les coûts de matériel, s'élevaient à 84 769 000 forints hongrois (Ft), soit environ 326 000 €.

104. En 2000, l'effet de conception a été d'environ 2 pour le revenu net disponible, 2,5 pour les dépenses alimentaires et 2 pour les dépenses totales.

105. Les résultats de l'enquête sont publiés annuellement en version bilingue (hongrois et anglais) avec une brève analyse des données. La dernière publication contenant les données de l'enquête HBS de 2001 a paru en 2002 sous le titre *House-*

hold Budget Survey: 2001 Annual Report (CSO, Budapest, 2002), Cette publication est également disponible sur CD-Rom.

Enquête sur la main-d'œuvre

106. L'enquête LFS est un nouveau type d'enquête sur les ménages lancé par le CSO en 1992. Son échantillon a été sélectionné en 1991 avec le recensement de 1990 comme cadre. Les villes se représentant elles-mêmes ont été définies comme celles qui comptaient 15 000 habitants ou plus. L'échantillon initial d'un trimestre comprenait 9 960 DE de 670 localités avec trois adresses par DE, ce qui a donné un échantillon trimestriel de $9\,960 \times 3 = 29\,880$ adresses.

107. Dans la seconde moitié des années 1990, on a assisté à une demande de données régionales plus détaillées de l'enquête LFS, et la taille de l'échantillon a été agrandie de 40 %. Le nombre de localités couvertes par l'échantillon, en particulier le nombre de DE, a également augmenté. En 2000, l'échantillon comptait 12 829 DE appartenant à 754 localités, de sorte que près de 36 500 ménages ont été visités chaque trimestre. Pour plus de détails sur cet échantillon agrandi, voir Éltető (2000).

108. Actuellement, la collecte de données a lieu chaque mois, avec la semaine du douze du mois comme période de référence et la semaine suivante comme période de collecte des données. Les données de l'enquête LFS sont recueillies principalement par des entrevues face à face à l'aide de traditionnels questionnaires sur papier, quoique l'on envisage de recourir de plus en plus à des entretiens téléphoniques, surtout pour les entretiens répétés. Aux adresses sélectionnées, toutes les personnes âgées de 15 à 74 ans peuvent être et sont interrogées pour l'enquête.

109. D'après le système de renouvellement appliqué à l'enquête LFS, les ménages choisis restent dans l'échantillon durant six trimestres consécutifs, puis le quittent. Cela veut dire qu'à chaque trimestre un sixième de l'échantillon est renouvelé.

110. Les pondérations de l'échantillon LFS sont calculées de la même manière que pour l'enquête HBS. Les pondérations finales de l'échantillon sont également déterminées selon l'approche du ratio de ratissage. Pour le calibrage de ces pondérations, on utilise les variables auxiliaires suivantes dans les 19 comtés ainsi que dans la capitale :

- Âge-sexe (2×10 catégories);
- Résidence dans des villes ayant les droits de comté ou ailleurs (2 catégories).

111. Les erreurs d'échantillonnage pour les données trimestrielles de l'enquête LFS figurant sur le « tableau principal » sont appliquées selon le modèle « jackknife » stratifié à l'aide du logiciel VPLX. Les erreurs d'échantillonnage pour les données mensuelles sont également calculées, mais elles ne sont pas publiées. Du point de vue de l'erreur d'échantillonnage, l'enquête LFS est conforme aux exigences de précision d'Eurostat énoncées dans la Règle appelée « Council Regulation » (EC), n° 577/98 du 9 mars 1998.

112. Les taux de non-réponse à l'enquête LFS, en particulier les taux de refus, sont beaucoup plus faibles que ceux enregistrés pour l'enquête HBS. Du début jusqu'à 1997, les taux de non-réponse ont progressé légèrement pour atteindre un maximum de 14,3 %. Après 1997, le taux total de non-réponse a diminué, tombant à 9,2 % en 2000. Les taux de refus ont également augmenté dans un premier temps, atteignant 7 % en 1996 et 1997, puis ont diminué, tombant à 3,2 % en 2000.

113. L'enquête LFS est coûteuse. En 2000, les dépenses directes se sont élevées à 109 802 000 forints, soit 422 000 €. Toutefois, si l'on considère qu'un même mé-

nage est contacté quatre fois par an et qu'aucune incitation à coopérer n'est donnée aux ménages, les dépenses par ménage sont nettement moins élevées que pour l'enquête HBS.

114. Principalement parce que l'échantillon LFS contient beaucoup plus d'UPE que l'échantillon HBS, l'effet de conception est considérablement plus faible. En 2001, l'effet de conception pour la taux de chômage a été de 1,4, tandis que pour le taux de participation féminine, cet effet a été de 0,8.

115. L'enquête LFS est complétée par un module axé sur certains thèmes tels que la situation des femmes qui travaillent, des questions concernant les mères en congé pour élever leurs enfants, etc. Ces modules sont inclus en moyenne pour trois des quatre trimestres. L'un de ces trois modules, en général celui du deuxième trimestre de l'année, porte sur le thème recommandé par Eurostat pour cette année. Les questionnaires de base de l'enquête LFS, de même que ceux qui sont ajoutés pour les modules Eurostat, contiennent toutes les informations exigées par Eurostat.

116. Les données trimestrielles comme les données annuelles de l'enquête LFS sont publiées dans des bulletins bilingues.

117. On peut conclure que les enquêtes HBS et LFS sont toutes deux de très importantes enquêtes sur les ménages pour le CSO hongrois. Les données de l'enquête HBS sont utilisées non seulement pour calculer les pondérations pour l'indice des prix à la consommation, mais aussi pour établir les valeurs trimestrielles et annuelles du produit intérieur brut (PIB). De plus, ces données sont d'une importance vitale pour certains domaines, tels que la recherche sur les conditions de vie des différentes couches sociales, les profils de dépenses de divers types de ménages et les changements constatés dans ces profils, la demande de différents types de biens de la part des consommateurs, etc. Il faut en outre préciser que, pour accentuer le respect des règles d'Eurostat, les dépenses de 2001 ont été groupées selon le système de Classification de la consommation individuelle par fonction (COICOP) [Nations Unies, 2000, troisième partie).

118. Bien que l'on connaisse d'autres sources le nombre de chômeurs déclarés, les données de l'enquête LFS s'écartent de ce nombre tant dans leur définition que dans leurs détails. Les informations sur la situation réelle et sur les changements observés sur le marché du travail telles qu'elles ressortent de l'enquête LFS sont indispensables aussi bien pour le gouvernement central que pour les administrations locales, ainsi que pour les chercheurs. Le taux de chômage officiel fondé sur les données LFS est l'un des indicateurs économiques les plus importants.

3. Conception et exécution des enquêtes sur les ménages en Lettonie⁴

Enquête sur le budget des ménages en Lettonie

119. L'enquête sur le budget des ménages (HBS) est une enquête continue qui a commencé en 1995. Elle a été restructurée en mai 2001.

120. L'enquête HBS a été lancée en septembre 1995 avec une assistance technique de la Banque mondiale. Elle en était déjà au stade préparatoire lorsqu'il a été exigé que ses résultats soient conformes aux normes d'Eurostat.

Portée de l'enquête

121. La population visée par l'enquête HBS est l'ensemble des ménages de Lettonie. Les personnes vivant dans des institutions (foyers pour personnes âgées, mai-

⁴ Préparé par Janis Lapins, Département de statistique, Banque de Lettonie ; Edmunds Vaskis, Zaiga Priede, Bureau central de statistique de Lettonie; et Signe Balina, Université de Lettonie, Riga.

sons pour enfants handicapés, foyers d'étudiants, hôtels, casernes, hôpitaux, sanatoriums, établissements pénitentiaires, etc.) et les sans-logis sont exclus de l'enquête actuelle.

Échantillonnage

122. L'échantillon représente l'ensemble de la population ainsi que ses groupes les plus typiques. Tous les mois, 342 ménages sont interrogés. Chaque ménage inclus dans l'échantillon est interrogé une seule fois.

123. On applique l'échantillonnage probabiliste stratifié à deux phases. Les ménages sont stratifiés selon leur degré d'urbanisation et leur localisation géographique. L'allocation de l'échantillon entre les strates se fait proportionnellement à la taille de la population des strates. Dans les zones urbaines, on a choisi le registre de population comme cadre d'échantillonnage, et dans les zones rurales, les listes de ménages.

124. Les six districts administratifs de Riga, la capitale, et les six autres grandes villes forment 12 strates à autoreprésentation. Toutes les autres villes sont utilisées comme UPE dans les autres zones urbaines, et sont réparties entre 10 strates que l'on a définies en combinant 5 régions et deux groupes différenciés par leur taille. A la première phase de l'échantillonnage, des UPE sont sélectionnées dans chaque strate avec des probabilités proportionnelles à leur nombre total d'habitants. A la seconde phase, on choisit des personnes âgées de 15 ans et plus par échantillonnage aléatoire.

125. Dans les zones rurales, les ménages sont répartis entre cinq strates ou régions géographiques. Il est de règle que les *pagasts* (communes civiles, les plus petites subdivisions administratives rurales) soient utilisées comme UPE; certaines des petites *pagasts* sont ajoutées à un territoire voisin. À l'intérieur de chaque strate, les UPE sont sélectionnées selon des probabilités proportionnelles à leur nombre de ménages. À la seconde phase, les ménages sont sélectionnés par échantillonnage aléatoire simple.

Coût de l'enquête

126. L'enquête HBS est l'un des exercices statistiques les plus coûteux. En 2001, le coût par ménage de l'enquête HBS a été de 24 lats lettons (LVL) [soit environ 40 \$]. Le principal élément de dépense était lié au travail de terrain. La rémunération des enquêteurs représentait 44 % du coût total de l'enquête, suivie des incitations aux répondants (26 %), des salaires des superviseurs (14 %) et des transports (8 %).

Erreur d'échantillonnage

127. Dans l'enquête HBS, les variances des estimations sélectionnées pour les principaux domaines étudiés (la capitale et les six autres grandes villes, d'autres villes et des zones rurales) sont estimées à l'aide du logiciel SUDAAN. Sur la base de ces estimations, on estime les variances et les effets de conception à l'échelle du pays.

Non-réponse

128. En 2000, le niveau total de non-réponse a été de 26,1. Les principales raisons de la non-réponse étaient le refus, y compris des ménages qui ont cessé de participer durant le mois de l'enquête (46,0 % de tous les cas de non-réponse), suivie de l'absence (31,8 %) et de l'incapacité de participer pour cause de maladie ou de vieillesse (11,6 %). Le taux de non-réponse était beaucoup plus élevé dans les zones urbaines (31,9 %) que dans les zones rurales (12,2 %).

129. Les ménages qui refusent de participer à l'enquête ou qui ne répondent pas aux questions qui leur sont posées, ainsi que les ménages ne se trouvant pas à

l'adresse donnée peuvent avoir un impact sur la précision des résultats, qui ne saurait être négligé. Pour maintenir la taille de l'échantillon au niveau choisi, on a appliqué l'approche de l'échantillonnage séquentiel. Le ménage qui refusait de répondre ou qui ne répondait pas a été remplacé par un autre tiré de la liste de réserve.

Restructuration de l'enquête HBS, 2001-2002

130. La dernière restructuration de l'échantillon HBS était fondée sur le recensement de la population effectué au printemps de 2000. Les instruments de l'enquête ont été profondément modifiés et la période de référence rétrospective de 12 mois a été introduite pour les biens durables, les achats et les paiements rarement effectués, le revenu saisonnier du travail rémunéré et les revenus et dépenses en espèces liés à la production agricole du ménage. L'enquête HBS précédente a cessé à la fin de 2000.

131. À partir de janvier 2002, les échantillons des enquêtes HBS et LFS ont été coordonnés. Pour ces deux enquêtes, l'échantillon annuel de ménages est réparti de façon homogène dans le temps (le même nombre de ménages participe à l'enquête durant chacune des 52 semaines de l'année). L'échantillon d'UPE est également réparti chaque trimestre de façon homogène sur l'ensemble du territoire.

132. Pour la nouvelle enquête HBS et l'enquête LFS continue, on utilise le même réseau d'enquêteurs. En revanche, on utilise des réseaux d'enquêteurs différents pour les anciennes enquêtes HBS et LFS. De plus, les enquêteurs opérant dans les zones rurales sont recrutés parmi la population locale. Selon la nouvelle formule, les enquêteurs sont mobiles et peuvent travailler sur différents territoires administratifs. Cela permet de répartir plus largement l'échantillon sur les territoires ruraux. (Dans la nouvelle enquête HBS, l'échantillon annuel est réparti sur 208 UPE rurales différentes.) En même temps, la charge de travail des enquêteurs est distribuée plus équitablement, et les transports sont assurés de façon plus économique. La nouvelle structure du Bureau central de statistique (BCS) a été instituée en janvier 2002.

L'enquête sur la main-d'œuvre en Lettonie

133. Au cours de la période 1995-2001, l'enquête sur la main-d'œuvre en Lettonie a été effectuée deux fois par an, en mai et en novembre. L'enquête LFS continue a commencé en 2002.

134. L'enquête LFS a été préparée conformément à la méthodologie de l'Organisation internationale du travail (OIT), approuvée à l'échelle internationale pour ce type d'enquêtes, qui permet les comparaisons avec d'autres pays (Eurostat, 1998a; 1998b).

Portée de l'enquête

135. La population soumise à l'enquête LFS comprend toutes les personnes âgées de 15 ans et plus résidant chez des ménages particuliers du pays. Les personnes vivant dans des institutions telles que foyers pour personnes âgées, maisons pour enfants handicapés, hôtels, casernes, hôpitaux, sanatoriums, établissements pénitentiaires, etc., ainsi que les sans-logis, sont exclus de l'enquête.

136. Conformément aux recommandations d'Eurostat et pour réduire les coûts de l'enquête, toutes les personnes de ce groupe d'âge qui vivent dans la même famille que les personnes sélectionnées sont également interrogées. L'échantillon national d'un cycle de l'enquête comprend 7 940 ménages.

137. Toutes les questions de l'enquête se rapportent à la semaine (du lundi au dimanche) qui précède le jour de l'entrevue. Normalement, les données sont re-

cueillies sur papier lors d'entrevues face à face. Si un répondant ne veut pas ouvrir sa porte, l'entrevue se fait au téléphone.

Échantillonnage

138. Pour les zones urbaines, l'échantillon est tiré du registre de population. Pour les zones rurales, il repose sur les listes complètes des ménages. Depuis 1998, l'échantillon rural est fondé sur le registre des ménages établi par le Bureau central de statistique de Lettonie.

139. L'enquête LFS porte sur 7 grandes villes, 32 villes plus petites et tous les *pagasts*. À chaque cycle, près de 16 000 personnes sont interrogées. Pour la construction de l'échantillon, la procédure suivie est celle de l'échantillonnage à une phase (dans les grandes villes et les zones rurales) ou à deux phases (dans les villes plus petites) avec stratification fondée sur le découpage administratif du pays. Dans les zones urbaines, un échantillon aléatoire simple de personnes âgées de 15 ans ou plus est sélectionné dans chaque UPE. Dans les zones rurales, un échantillon aléatoire simple de ménages est sélectionné dans chaque *pagast*.

140. D'après le plan de renouvellement de l'échantillon de l'enquête LFS, des membres de chaque ménage sont inclus à trois reprises dans l'enquête. À chaque cycle, le taux de renouvellement de l'échantillon est d'un tiers des ménages de chaque grande ville, ville plus petite ou *pagast*.

Non-réponse

141. En 2000, le taux de non-réponse a été de 10,1 %. Il a été plus faible dans les zones rurales (8,5 % seulement) que dans les zones urbaines (11,4 %). Dans les zones rurales, le pourcentage de refus a été particulièrement modeste, d'environ 0,5 % seulement. Des entrevues de substitution ont été utilisées pour accroître le taux de réponse. Environ un tiers des entrevues ont été effectuées avec des répondants de substitution.

Imperfections du cadre

142. Toutes les personnes sélectionnées n'habitaient pas à l'adresse indiquée dans le registre comme celle de leur logement. Comme la recherche et l'interrogation de ces personnes sont coûteuses, prennent du temps et s'avèrent parfois impossibles, les enquêteurs se voient contraints d'interroger les ménages qui vivent effectivement à l'adresse sélectionnée. L'analyse des cas de non-participation a montré que 2,0 % seulement des cas de non-participation (2,3 % dans les zones rurales) étaient identifiés comme étant liés à des imperfections du cadre (logement vide, maison démolie, adresse inexistante, etc.).

Restructuration de l'enquête LFS, 2001-2002

143. Le questionnaire de l'enquête LFS a été restructuré en 2001 de manière à répondre pleinement aux exigences de l'Union européenne (UE). À présent, cette enquête est continue.

144. Depuis janvier 2002, d'importantes modifications ont été également apportées à la conception de l'échantillon de l'enquête LFS. Le même réseau d'enquêteurs est utilisé pour les enquêtes LFS et HBS. De ce fait, à partir de janvier 2002, les échantillons utilisés pour les deux enquêtes ont été coordonnés. On compte que cette coordination favorisera une utilisation plus efficace des ressources affectées à ces enquêtes.

145. Les enquêteurs ont reçu une formation en décembre 2001. L'enquête LFS continue à démarré en janvier 2002.

Autres enquêtes sur les ménages

146. L'enquête sur les conditions de vie [*Living Conditions Survey* (LCS)] a été lancée en 1994 et en 1999 dans le cadre du projet NORBALT financé par le Gouvernement norvégien et exécuté en étroite coopération avec l'Institut Fafo (Institut de sciences sociales appliquées d'Oslo).

147. Plusieurs autres enquêtes sur les ménages ont été lancées dans la deuxième moitié des années 1990, notamment l'Enquête sur la famille et la fécondité (1995), l'Enquête sur l'emploi du temps (1996), l'Enquête sur la confiance des consommateurs (1993-1999), l'Enquête sur la consommation d'énergie des ménages (1996), l'Enquête sur le tourisme national (1998), l'Enquête sur l'utilisation d'ordinateurs individuels par les ménages (1998), le Module spécial d'enquête sur la pauvreté (1998), l'Enquête sur les attitudes à l'égard des problèmes de suicide (1999), etc.

148. Depuis 1996, l'Enquête sur les voyageurs à la frontière a lieu trois ou quatre fois par an. Cette enquête recense à la fois les flux de voyageurs lettons de retour de l'étranger, et les flux de voyageurs étrangers quittant la Lettonie.

149. Il est de règle que les résultats des enquêtes soient publiés en letton et en anglais et soient disponibles sur papier ou sous forme électronique. Pour les besoins de la recherche, le BCS assure l'accès aux fichiers de microdonnées aux utilisateurs lettons et étrangers.

Quelques conclusions

150. Nous comptons que la constitution d'un nouveau groupe d'enquêteurs hautement spécialisés et mobiles donnera plus de souplesse à la planification et à l'exécution des nouvelles enquêtes par sondage et de différentes enquêtes spéciales.

151. Le BCS envisage également d'adopter les méthodologies modernes de collecte de données. L'une des premières mesures à cet égard sera la mise en place dans les prochaines années de la technologie d'entrevues personnelles assistées par ordinateur [*computer-assisted personal interview* (CAPI)].

4. Les enquêtes par sondage sur les ménages en Lituanie⁵

Introduction

152. L'Enquête sur le budget des ménages (HBS) a été la première enquête par sondage effectuée en Lituanie. Elle a eu lieu pour la première fois sur une période de 12 mois en 1936-1937. Cette enquête a été la seule enquête par sondage utilisée pour produire des statistiques pour l'économie planifiée de la Lituanie. Une fois que le pays a accédé à l'indépendance, en 1990, l'économie nationale s'est tournée vers l'économie de marché. Un nouveau questionnaire a dû être conçu pour recueillir davantage de données, et il fallait un échantillon d'une conception nouvelle pour couvrir le secteur privé; par ailleurs, les résultats publiés ont dû être redéfinis de manière à permettre aux utilisateurs de disposer de données comparables à celles des autres pays. Cette refonte de l'enquête HBS s'est faite en 1996 avec l'aide d'experts de la Banque mondiale et une description en est donnée dans Šniukstiene, Vanagaite et Binkaustiene (1996). La conception de l'échantillon et la méthode d'estimation restent les mêmes.

153. Les autres enquêtes habituelles sur les ménages sont l'enquête sur la main-d'œuvre (LFS), qui a débuté en 1994. Le registre de population a été modernisé en 1996. Depuis lors, il sert de cadre d'échantillonnage pour la sélection de l'échantillon pour la plupart des enquêtes sur les ménages, y compris l'enquête LFS.

⁵ Préparé par Danute Krapavickaitė, Institut de mathématiques et d'informatique, 4 Akademijos str., LT 2600 Vilnius et Département de statistique de Lituanie, 29 Gedimino Avenue, 2746 Vilnius.

154. D'autres enquêtes sur les ménages, essentiellement ponctuelles, ont porté sur certains thèmes tels que les conditions de vie (1997), l'emploi du temps (1998), la consommation d'énergie des ménages (1997), l'accès aux services de santé (1998) et l'utilisation d'ordinateurs par les ménages (2000).

Estimations et erreurs dans l'enquête sur la main-d'œuvre

Conception de l'échantillon

155. La population visée par l'enquête LFS comprend les habitants de la Lituanie âgés de 15 ans ou plus. L'échantillon est construit de la façon suivante : après avoir sélectionné un échantillon aléatoire simple d'environ 3 000 personnes sur le registre de population, on ajoute les membres de leurs ménages à l'échantillon, même s'ils ne figurent pas sur le registre. Les femmes représentent 52,5 % des répondants.

Renouvellement de l'échantillon

156. Pour éviter de profonds changements dans les résultats d'un cycle de l'enquête à un autre, on ne renouvelle qu'un tiers de l'échantillon à chaque cycle. Chaque ménage sélectionné participe à deux cycles; il est retiré pour un cycle, réintroduit pour le suivant, puis est éliminé du système.

Estimations et leur précision

157. La distribution des répondants entre zones urbaines et rurales, par âge et par sexe, diffère légèrement des distributions correspondantes fondées sur les données du recensement. La post-stratification de l'échantillon se fait selon 12 groupes d'âge, les deux sexes et 10 comtés, donnant un total de 240 groupes pondérés.

158. Différents systèmes de pondération sont utilisés pour l'estimation des personnes ayant un emploi et des chômeurs. Pour améliorer la précision des estimations des chômeurs, on utilise également les indices de la bourse du travail pour la post-stratification. Les estimations de la variance sont décrites dans Krapavickaitė, Klimavicius et Plikusas (1997) pour la conception d'un échantillon de taille fixe.

Coût de l'enquête

159. Le coût d'une enquête est d'environ 70 000 litai lituaniens⁶. L'impression des questionnaires et l'expédition aux répondants représentent 14 % de ce coût total, et les 86 % restants servent à couvrir la rémunération des enquêteurs, les frais de transport des enquêteurs et les coûts d'envoi par la poste des questionnaires remplis à Vilnius. Les dépenses afférentes au travail méthodologique de conception de l'échantillon, de préparation des questionnaires, de sélection de l'échantillon et d'entrée, d'édition et de traitement des données ne sont pas incluses.

Enquête sur le budget des ménages

Conception de l'échantillon

160. L'enquête HBS est une enquête continue. L'échantillon est sélectionné une fois par an et divisé en 12 parties qui sont distribuées une sur chaque mois. Chaque ménage sélectionné participe à l'enquête pendant un mois. La population des ménages lituaniens est divisée en trois strates, selon leur type de résidence. Un échantillon aléatoire simple de 4 476 personnes âgées de 16 ans ou plus est sélectionné à partir du registre de population des plus grandes villes : Vilnius, Kaunas, Klaipėda, Šiauliai et Panevėžys. Un échantillon aléatoire de 20 grappes avec des probabilités proportion-

⁶ Taux de change (US dollars de 2000); 1 US dollar = 4 litai.

nelles à leur taille est prélevé sur les 140 grappes des petites villes, et un échantillon aléatoire de 33 grappes avec des probabilités proportionnelles à leur taille est prélevé sur la population de 463 grappes à la première phase. Un échantillon aléatoire simple de personnes est prélevé sur chaque grappe sélectionnée. Toutes les personnes vivant dans les ménages sélectionnés sont interrogées. Lorsque plusieurs ménages sont à la même adresse sélectionnée, on inclut dans l'échantillon le ménage auquel appartient la personne dont l'anniversaire est le plus proche.

Les estimations et leur précision

161. Des pondérations sont utilisées pour l'estimation de l'enquête HBS. Les effets de conception des estimations sont supérieurs à l'unité, ce qui donne à penser qu'il conviendrait d'utiliser des informations auxiliaires à l'avenir pour obtenir des estimations plus précises.

Coût de l'enquête

162. Le coût annuel total de l'enquête est d'environ 900 000 litai, dont 61 % pour la rémunération des enquêteurs, 18 % d'impôts, 14 % de paiement aux ménages, 5 % de frais de transport et 2 % de frais divers.

Diffusion des résultats

163. Les résultats des enquêtes sont publiés par Statistics Lithuania. Les principaux résultats sont publiés dans la revue mensuelle *Economic and Social Development in Lithuania*. Tous les résultats sont publiés dans des éditions spéciales sur des thèmes tels que la main-d'œuvre, l'emploi et le chômage et le revenu et les dépenses des ménages.

Conclusions

164. Les résultats provisoires du Recensement de la population et des logements de 2001 estiment la population totale de la Lituanie à 3 491 000 habitants. Ce chiffre est inférieur de 202 000 habitants à celui tiré des données démographiques publiées le 1^{er} janvier 2001. Après avoir finalisé les résultats du recensement, Statistics Lithuania disposera de données démographiques plus fiables à partir desquelles il lui sera possible d'améliorer les enquêtes à venir sur les ménages. L'erreur systématique devrait diminuer pour ces enquêtes.

5. Les enquêtes sur les ménages en Pologne pendant la période de transition⁷

Introduction

165. En Pologne, les enquêtes sur les ménages sont le produit d'une assez longue tradition [Główny Urząd Statystyczny (Office central de statistique de Pologne, (GUS)], 1987, 1998a, 1999; Kordos, 1985, 1996; Lednicki, 1982). Dans les années 1980, le Système intégré d'enquêtes sur les ménages [*Integrated System of Household Surveys (ISHS)*] a été progressivement mis en place. Lancé en 1982, il est devenu pleinement opérationnel en 1992 (GUS, 1987; Kordos, 1985).

166. Le principal élément de l'ISHS était l'enquête sur le budget des ménages (HBS), qui était fondé sur un échantillonnage à deux phases, un renouvellement trimestriel des ménages sur une année et un renouvellement des ménages par tiers sur les trois années suivantes. Cela veut dire que deux tiers des ménages étaient inclus dans le panel pendant quatre années consécutives. Il y avait aussi un cycle de quatre ans pour

⁷ Préparé par Jan Kordos, de la Faculté d'économie de Varsovie; et par Bronislaw Lednicki et Malgorzata Zyra, de l'Office central de statistique, Al. Niepodleglosci 208, 00-925, Varsovie.

l'étude de sous-échantillons. Il a été mis fin à ce programme d'enquêtes en 1992. En même temps, c'est-à-dire durant les années 1983-1992, les sous-échantillons sélectionnés pour l'enquête HBS ont été utilisés pour plus de 30 enquêtes sociales utilisant des modules par thème.

167. Les tentatives d'intégration des enquêtes sur les ménages effectuées dans les années 1980 ont grandement facilité l'ajustement des enquêtes sur les ménages sur les normes européennes (GUS, 1997). La méthodologie utilisée pour ces enquêtes a besoin d'une plus grande intégration et de nouvelles améliorations (Kordos, 1998),

Les enquêtes sur les ménages dans la période de transition

168. Après 1990, ces enquêtes ont été considérablement étendues et modifiées. L'enquête HBS est encore en cours d'amélioration, et en 1992, une nouvelle enquête LFS a été lancée. Par ailleurs, d'autres enquêtes ont été effectuées, y compris une enquête sur les conditions de vie, une enquête sur l'état de santé des ménages, une enquête sur l'emploi du temps, un microrecensement de la population et diverses enquêtes post-énumération.

L'enquête sur le budget des ménages

169. Les enquêtes sur le budget des ménages obéissent à une tradition qui remonte à près de 45 ans (GUS, 1999; Kordos, 1996; Lednicki, 1982). Diverses méthodes d'enquête ont été expérimentées et des essais ont été tentés pour améliorer leur exécution. Au début des années 1990, la méthodologie de l'enquête a été changée. Dans la nouvelle méthode d'exécution de l'enquête HBS mise en place en 1992, la classification des revenus et des dépenses ainsi que celle des types socio-économiques ont été remplacées. Pour la première fois, tous les types de ménages polonais, regroupés dans un échantillon d'environ 32 000 ménages, ont été inclus dans l'enquête. En 1997, on s'est efforcé d'améliorer l'intégration des enquêtes sur les ménages⁸. En 2000, l'enquête HBS a été restructurée et certains éléments méthodologiques ont été modifiés (Kordos, Lednicki et Zyra, 2002). D'autres améliorations de l'enquête HBS et son intégration à d'autres enquêtes sur les ménages sont envisagées, en grande partie sur les recommandations de l'Eurostat (Eurostat, 1997).

⁸ Voir règle interne n° 20 du Président de l'Office central de statistique du 30 octobre 1997 sur la création du Groupe de travail pour l'amélioration de la méthodologie et l'intégration des enquêtes sur les ménages.

L'enquête sur la main-d'œuvre

170. L'enquête sur l'activité économique de la population a été lancée pour la première fois en Pologne en mai 1992, et s'est répétée chaque trimestre jusqu'au troisième trimestre de 1999 (Szarkowski et Witkowski, 1994). Elle a été conçue selon les recommandations de l'OIT. Chaque trimestre, quelque 24 000 ménages et des membres de ces ménages âgés de 15 ans ou plus ont été interrogés. À l'occasion, des modules sur certains thèmes sociaux ont été inclus dans l'enquête, augmentant sensiblement la possibilité d'analyses sociales et économiques ainsi que le champ des résultats publiés.

171. Les résultats de cette enquête sont publiés chaque trimestre. L'enquête LFS a été restructurée en 1999 de manière à être adaptée au nouveau découpage administratif du pays et à améliorer son efficacité conformément aux exigences d'Eurostat (Eurostat, 1998b; Verma, 1995).

Le microrecensement de la population et des logements de 1995

172. Plusieurs enquêtes spéciales ont été effectuées au cours de la dernière décennie, dont la plus grande a été le Microrecensement de 1995. En mai 1995, une vaste

enquête par sondage (microrecensement de la population et des logements) a été effectuée (Bracha, 1996; GUS, 1998a). Il s'agissait du troisième microrecensement; les deux précédents avaient eu lieu en 1974 et en 1984. Ces recensements sont l'occasion de saisir des données sur les personnes handicapées, les migrations et d'autres questions sociales.

173. Le microrecensement de 1995 portait sur 5 % de la population, c'est-à-dire près de 600 000 ménages. Le recensement complet de la population et des logements a été effectué en mai et juin 2002; le précédent avait eu lieu en 1988.

Enquêtes sur les conditions de vie

174. Outre l'enquête HBS, en 1997, il a été décidé d'effectuer une enquête sur plusieurs aspects des conditions de vie de la population (Kordos, Lednicki et Zyra, 2002). Cette enquête a été préparée avec soin, en collaboration avec des experts de l'Institut national de la statistique et des études économiques (INSEE) [France] et effectuée pour la première fois au milieu de 1997 sur un vaste échantillon. Elle a été reprise ensuite chaque année dans des proportions plus modestes sur des sous-échantillons, puis de nouveau à grande échelle toutes les quelques années.

175. Au total, 12 524 ménages ont participé à la première enquête, et le taux de réponse a été respectivement de 87 % pour les ménages et de 86 % pour les adultes. Au milieu de 1988, l'opération a été répétée à une échelle plus modeste.

176. En 1999, l'échantillon était composé de deux sous-échantillons : le sous-échantillon sélectionné en 1998 (panel) et un nouveau sous-échantillon de même taille que le panel de 1998. De cette façon, il y a eu chaque année un panel et un nouveau sous-échantillon sélectionné à partir du cadre d'échantillonnage actualisé.

177. Une nouvelle enquête de grande ampleur sur les conditions de vie a été lancée en 2001, sur un échantillon d'environ 24 000 ménages, sur lesquels 18 052 ont répondu, ce qui représente un taux de non-réponse de 25 %. Cette enquête doit se poursuivre jusqu'au lancement, en 2005, d'une nouvelle enquête sur les revenus et les conditions de vie [*Income and Living Conditions Survey* (EU-SILC)] préparée conformément au programme d'Eurostat (Eurostat, 2001).

Enquête sur l'état de santé de la population

178. Cette enquête a été effectuée en avril 1996 sur 192 000 ménages. Le taux de non-réponse a été de 88,6 %. Il s'agissait de la première enquête effectuée à une aussi grande échelle en Pologne sur l'état de santé de la population.

179. Cette enquête était conforme aux recommandations de l'Organisation mondiale de la santé (OMS) qui permettent de comparer les résultats obtenus avec ceux d'autres pays européens, en particulier les États membres de l'Union européenne et les pays de la Commission économique pour l'Europe (CEE).

Enquête sur l'emploi du temps

180. Le GUS a effectué des enquêtes sur l'emploi du temps en 1969, 1976 et 1984 (Kordos, 1988b). En 1996, le GUS a effectué une enquête plus modeste, sur un échantillon de 1 000 ménages, dans laquelle ont été inclus tous les membres de ces ménages âgés de 10 ans ou plus. L'un des objectifs de cette enquête était notamment de vérifier la possibilité d'application de la méthodologie proposée par Eurostat (GUS, 1998b). Une enquête à grande échelle sur l'emploi du temps doit être menée en 2004.

**Aspects méthodologiques communs
à toutes les enquêtes sur les ménages**

Cadres d'échantillonnage

181. Les recensements de la population servent de base aux cadres d'échantillonnage utilisés dans les enquêtes sur les ménages en Pologne. Les unités primaires d'échantillonnage (UPE) sont formées sur la base des districts statistiques d'énumération (DSE) ou des zones d'énumération du recensement (ZEC) généralement ajustées en fonction des exigences spécifiques de l'enquête. Les logements servent généralement d'unités secondaires d'échantillonnage (USE). La liste des logements des DSE ou des ZEC est actualisée chaque année, et cette actualisation englobe à la fois un accroissement du nombre des logements, par suite de la construction de nouveaux immeubles, et une diminution de ce nombre dû à la démolition, ainsi que des rectifications du tracé des limites des districts par suite de changements dans le découpage administratif du pays. Pour chaque district, le cadre d'échantillonnage contient des données sur les adresses et une estimation de la population et du nombre de logements (GUS, 1998a).

182. Pour la sélection de l'échantillon à utiliser pour les enquêtes HBS et LFS, il a fallu combiner des DSE ou des ZEC de manière à parvenir à la taille minimum requise pour chaque UPE. C'est ainsi que 29 172 UPE ont été formées pour l'enquête HBS à partir de 33 023 DSE (dans les zones urbaines, les UPE devaient compter au moins 250 logements, et dans les zones rurales, 150).

La conception des échantillons pour les enquêtes sur les ménages

183. Généralement, pour chaque enquête sur les ménages, on sélectionne l'échantillon en deux phases avec une probabilité proportionnelle à la taille (PPT). La stratification est fondée sur la région (voivodship), les zones rurales et urbaines et, dans certains cas, la taille de la localité. Pour les enquêtes continues, telles que les enquêtes HBS et LFS, un différent mode de rotation a été utilisé, et les résultats finals sont pondérés pour limiter l'impact des cas de non-réponse.

Conception de l'échantillon pour l'enquête HBS

184. Au cours des 45 dernières années, différents types de modèle d'échantillon ont été utilisés (GUS, 1999; Kordos, 1996; Lednicki, 1982). Ici, nous examinons le modèle le plus récent, qui est en place depuis 2000). Les modèles d'échantillon utilisés pendant la période 1992-2000 sont décrits de façon détaillée dans Kordos, Lednicki et Zyra (2002).

185. Depuis 2001, deux sous-échantillons de 675 UPE ont été choisis à partir d'un total de 29 172 UPE. Les UPE sont stratifiées selon 16 voivodships, et dans chaque voivodship, selon la catégorie de taille des localités. Les grandes villes forment une strate distincte. Chaque voivodship compte de 3 à 12 strates. Au total, il y a 96 strates. L'allocation de l'échantillon aux strates est proportionnelle à la population totale des logements de chaque strate. Les UPE sont sélectionnées selon une probabilité proportionnelle au nombre de logements, conformément à la formule de Hartley-Rao. Dans chaque UPE, 24 logements sont sélectionnés pour deux ans (deux logements pour chaque mois, et le même nombre de logements est interrogé chaque année). De plus, dans chaque UPE, 150 logements ont été sélectionnés de façon indépendante comme sous-échantillon de réserve à utiliser en cas de non-réponse. Chaque année, un nouveau sous-échantillon de 675 UPE sera choisi pour deux ans.

Pondération pour l'enquête HBS

186. Pour l'enquête HBS, les taux de non-réponse sont généralement élevés et influent considérablement sur la structure socio-économique des ménages de l'échantillon. Pour limiter cet impact, les résultats de l'échantillon sont pondérés.

187. D'abord, chaque ménage de l'échantillon est pondéré en proportion inverse de sa probabilité de sélection. On utilise des pondérations de sources extérieures. Pour l'enquête HBS, on applique d'autres pondérations appropriées en plus de celles de l'enquête LFS (pour la taille des ménages et les proportions entre populations urbaine et rurale).

Méthode d'estimation de l'erreur type

188. Jusqu'en 2000, la méthode d'estimation de l'erreur type a été celle du groupe choisi au hasard. Depuis 2001, on utilise une méthode de demi-échantillons équilibrés.

Conception de l'échantillon pour l'enquête LFS et nouvelle conception adoptée en 1999

189. Pour l'enquête LFS, on a choisi un échantillon en deux phases avec stratification. Les UPE étaient les ZEC dans les villes et les DES dans les zones rurales. (Dans certains cas, les unités d'échantillonnage ont été créées par la fusion de deux ZEC ou DES voisines ou plus pour répondre aux exigences de taille minimum.) Des logements ont servi d'unités secondaires d'échantillonnage (Szarkowski et Witkowski, 1994).

Refonte de l'enquête en 1999

190. Depuis le quatrième trimestre de 1999, l'enquête LFS est effectuée de façon continue. Les UPE et les USE ont été sélectionnées de la même manière que lors de l'enquête précédente, mais l'allocation de l'échantillon par 16 voivodships a changé. Pour obtenir une plus grande précision des estimations par voivodship, la taille de l'échantillon d'un voivodship a été allouée de façon presque proportionnelle à la racine carrée du nombre de logements du voivodship. Les tailles des strates créées à l'intérieur des voivodships étaient proportionnelles à la taille des localités.

191. À l'intérieur des strates, les UPE ont été sélectionnées selon une probabilité proportionnelle au nombre de leurs logements. Ensuite, un nombre déterminé de logements (de quatre à neuf) ont été sélectionnés pour chaque UPE. Toutes les 13 semaines de chaque trimestre⁹, les enquêteurs visitent un nombre déterminé de logements sélectionnés au hasard (1880-1900) et collectent des données concernant l'activité économique de leurs habitants au cours de la semaine précédente. L'enquête couvre toutes les personnes âgées de 15 ans ou plus qui vivent dans les logements sélectionnés. Un échantillon des logements à visiter est changé chaque semaine. Les échantillons hebdomadaires proviennent d'un découpage aléatoire de l'échantillon trimestriel en 13 fractions. L'échantillon trimestriel compte de 24 440 à 24 700 logements (GUS, 2000).

192. La participation des ménages à l'enquête se fait par rotation, à raison de deux trimestres de participation, deux trimestres de non-participation, puis à nouveau deux trimestres de participation, après quoi les ménages sont retirés du système [profil de rotation 2-(2)-2].

⁹ D'après les règlements d'Eurostat, le « trimestre » actuellement appliqué à l'enquête LFS diffère légèrement du trimestre du calendrier, en ce sens que chaque trimestre de l'enquête LFS est constitué de 13 semaines et commence toujours un lundi. Ainsi, le premier trimestre de 2000 est allé du 3 janvier au 3 avril.

Pondération des résultats de l'enquête LFS

193. La pondération se fait en trois phases (pour plus de détails, voir Kordos, Lednicki et Zyra, 2002).

Estimation des erreurs types

194. Jusqu'en 1999, les erreurs types des estimations se calculaient selon la méthode du groupe choisi au hasard. Depuis l'adoption de la nouvelle conception de l'enquête LFS au quatrième trimestre de 1999, c'est la technique de linéarisation de Taylor qui est utilisée.

Coûts des enquêtes sur les ménages

195. L'Office central de statistique de Pologne a son propre système d'estimation des coûts des enquêtes sur les ménages. Pour chaque enquête par sondage, le coût direct de l'enquête est estimé sur la base de l'expérience sur le terrain et de certaines recommandations administratives. Cette estimation englobe les coûts des entrevues sur le terrain, les frais de déplacement, les coûts de matériel, les coûts des services liés à l'enquête, les incitations données pour accroître la participation à l'enquête, les impôts, etc. (GUS, 2001). Parmi les coûts non inclus dans cette estimation figurent le codage et l'édition, l'utilisation d'ordinateurs, les coûts indirects et les frais généraux et le coût du personnel dont les responsabilités s'étendent à plusieurs projets.

196. À titre d'exemples, des éléments de coûts des enquêtes HBS et LFS en Pologne en 2000 ont été présentés à la section A du présent chapitre.

Effets de conception

197. Pour les enquêtes sur les ménages en Pologne, c'est-à-dire les enquêtes HBS et LFS, les effets de conception ont été calculés pour plusieurs caractéristiques (Kordos, Lednicki et Zyra, 2002). À titre d'exercice et pour comparer avec d'autres pays, on a calculé les effets de conception pour plusieurs paramètres pour les années 2000 et 2001.

198. Pour certaines caractéristiques de l'enquête HBS, les effets de conception et les erreurs types relatives (en pourcentage, données entre parenthèses) étaient les suivants : revenu total : 4,24 (1,1); dépenses totales : 4,16 (1,0); dépenses alimentaires : 3,53 (0,4); habillement et chaussures : 2,72 (1,5); entretien des logements : 4,04 (1,3); santé : 3,28 (1,7); transports et communications : 2,16 (4,5); et éducation : 2,50 (3,9).

199. Pour l'enquête LFS, en 2000 et 2001, les effets de conception ont été calculés pour le nombre total de chômeurs dans différents groupes à classification croisée établis sur la base de leur appartenance à des zones urbaines ou rurales, de la taille de leur localité (catégories de villes) et de leur niveau d'instruction. La plus forte dispersion a été observée pour les catégories de villes où les effets de conception se situaient entre 1,7 et 3,55.

200. Comme le montrent les estimations ci-dessus, les effets de conception pour les données des enquêtes HBS et LFS étaient généralement supérieurs à l'unité, et pour certaines caractéristiques, ils dépassaient même 4. Par conséquent, les erreurs types fondées sur les hypothèses d'échantillons aléatoires simples tendaient à sous-estimer les erreurs types découlant de l'application d'un échantillon de conception complexe.

Non-réponse dans les enquêtes sur les ménages

201. Comme on l'a vu à la section A, les taux de non-réponse ont augmenté dans les enquêtes HBS et LFS au cours de la dernière décennie. Les principales raisons de ces accroissements étaient les cas de refus et d'absence. Pour l'enquête HBS, les taux de refus sont passés de 10,2 % en 1992 à 25,0 % en 2000, et les taux d'absence, de 4,5 % en 1992 à 14,5 % en 2000.

202. En moyenne annuelle, les taux de non-réponse à l'enquête LFS ont augmenté régulièrement durant toute la période 1992-2000, passant de 4,5 % en 1992 à 22,1 % en 2000. Les taux de refus ont eux-même progressé de 2,0 % en 1992 à 10,9 % en 2000. Ce fort accroissement des taux de non-réponse observé durant la période 1992-2000 avait pour principales raisons les refus et les absences.

203. Les taux de non-réponse à l'enquête LFS diffèrent selon la taille des localités, les taux les plus élevés étant observés à Varsovie et les plus faibles dans les zones rurales. En 2000, les taux annuels moyens pondérés selon la taille des localités ont été les suivants : Varsovie : 54,5 %; grandes villes (entre 500 000 et un million d'habitants) : 32,6 %; villes moyennes (de 100 000 à 500 000 habitants) : 33,3 %; villes de 20 000 à 100 000 habitants : 23,1 %; villes de moins de 20 000 habitants : 19,0 %; et zones rurales : 11,1 %.

Conclusion

204. Dans cette section, ont été présentées des descriptions générales des enquêtes sur les ménages de l'Office central de statistique de Pologne (GUS) pendant la période de transition, et plus particulièrement de deux enquêtes continues, à savoir les enquêtes HBS et LFS. Le GUS a une longue tradition et une vaste expérience des enquêtes sur les ménages dans ces domaines. Cela a été utile au début de la période de transition pour la restructuration des enquêtes et la conception de nouvelles enquêtes.

205. L'assimilation des résultats du recensement de la population et des logements effectué en 2002 sera l'une des tâches les plus importantes des enquêtes sur les ménages dans les années à venir. Le recensement offrira non seulement des cadres d'échantillonnage actualisés pour les enquêtes sur les ménages mais aussi des informations auxiliaires qui permettront d'accroître la précision des estimations et qui pourront servir pour les méthodes d'estimation étroitement localisées actuellement à l'étude.

206. Nous avons commencé à préparer une nouvelle enquête sur les ménages – l'enquête sur les revenus et les conditions de vie (EU-SILC), qui doit être lancée en 2005 (Eurostat, 2001) – et nous améliorons les enquêtes en cours pour mieux les adapter aux normes de l'UE.

6. L'enquête sur la main-d'œuvre et l'enquête sur le budget des ménages en Slovénie¹⁰

Introduction

207. La République de Slovénie a accédé à l'indépendance au début des années 1990. Avant cela, intégrée à l'ex-Yougoslavie, ses activités statistiques étaient centralisées au sein de l'Office fédéral de statistique. À cette époque, les enquêtes sur les ménages n'occupaient pas une place importante dans le programme national de statistiques. Avec l'indépendance, l'Office de statistique slovène est rapidement passé de l'état de bureau régional à celui de bureau national. Le processus de transition s'est déroulé

¹⁰ Préparée par Vasja Vehovar, Faculté des sciences sociales, Université de Ljubljana ; et Metka Zaletel, Tatjana Novak, Marta Arnež et Katja Rutar, Office de statistique de la République de Slovénie.

sans heurts, en partie parce que la haute direction est restée la même pendant toute la période de transition.

L'enquête sur la main-d'œuvre (LFS)

Historique

208. La première enquête LFS a été effectuée en 1989 par la Faculté des sciences sociales de l'Université de Ljubljana (Vehovar, 1997). L'Office de statistique de la République de Slovénie a assumé l'entière responsabilité de l'enquête LFS en 1995.

209. Les échantillons de l'enquête LFS pour la période 1989-1995 ont été conçus et établis de façon assez pragmatique en raison des incertitudes du budget annuel. À partir de 1992, la conception a reposé sur un échantillon en grappes à trois phases comptant 3 000 nouveaux ménages chaque année. Les unités restaient dans l'échantillon pendant trois années consécutives, et l'échantillon total comptait environ 8 500 ménages.

Restructuration

210. En 1997, une profonde restructuration a eu lieu pour répondre à la demande de résultats plus fréquents (trimestriels) et plus détaillés (régionaux). Les directives d'Eurostat ont en outre contribué de façon importante à stimuler cette restructuration.

211. L'enquête LFS a été révisée et transformée en enquête continue par panel avec sélection trimestrielle d'un échantillon et publication trimestrielle des résultats. Chaque échantillon trimestriel est fractionné en six tranches de deux semaines. La période de référence pour les entrevues est la semaine (du lundi au dimanche) qui précède l'entrevue. On applique le système de rotation 3-1-2 : les ménages sont interrogés pendant trois trimestres consécutifs, puis omis pendant un trimestre, puis de nouveau inclus pour deux autres trimestres. Cette formule se traduit par un chevauchement de 60 % entre deux trimestres consécutifs et de 40 % seulement entre deux années de suite.

212. Le cadre d'échantillonnage de l'enquête LFS est le registre central de la population combiné à une stratification. Les définitions des strates sont fondées sur 6 types de peuplements (selon la taille des peuplements et la proportion d'agriculteurs parmi la population) et 12 régions géographiques. Après fractionnement, il y a un total de 47 strates.

213. Dans chaque strate, l'échantillon est sélectionné selon un échantillonnage systématique d'abord aléatoire. Une stratification implicite est appliquée par décomposition des données par peuplement, rue et numéro d'immeuble. Le taux d'échantillonnage de chaque strate est ajusté de manière à tenir compte du taux de non-réponse attendu. On n'a pas recours à des substitutions sur le terrain, car il a été démontré que ces substitutions présentent peu d'avantages et posent beaucoup de problèmes (Vehovar, 1999).

214. Chaque trimestre, 2 000 nouvelles unités sont sélectionnées. Par ailleurs, 5 000 ménages (répondants) sont ajoutés des quatre trimestres précédents. Ainsi, quelque 7 000 ménages sont sélectionnés chaque trimestre (2 000 du nouvel échantillon et 5 000 de la continuation de l'échantillon existant). Sur ce total, 6 000 devraient répondre. Le nombre total d'entrevues effectuées est d'environ 20 000.

Exécution

215. Tous les ménages du nouvel échantillon trimestriel sont interrogés personnellement (entrevues face à face) avec l'aide d'ordinateurs (CAPI). Quelque 30 enquêteurs expérimentés sont affectés à l'enquête LFS et sont tous équipés d'un ordinateur portable. Les entrevues suivantes se font depuis le central téléphonique de l'Office de statistique via CATI, sauf pour les ménages qui n'ont pas le téléphone et pour ceux qui ne peuvent participer à une entrevue téléphonique. Le taux de couverture téléphonique est d'environ 95 %. Avant l'entrevue, chaque ménage reçoit une lettre décrivant l'enquête et une brochure avec les résultats de précédentes enquêtes LFS. Il n'est pas offert d'incitations.

216. Pour les entrevues face à face de l'élément nouveau de l'échantillon, le taux de non-réponse a été de 17-18 % et les taux de refus, de 12-13 %. Pour les entrevues téléphoniques répétées avec les ménages appartenant déjà au panel, les taux de non-réponse sont légèrement plus faibles (10-11 %) et les taux de refus aussi (6-7 %). Le taux de non-réponse à l'enquête LFS a augmenté considérablement à partir de 1991, mais s'est stabilisé au cours des quatre dernières années.

Erreurs d'échantillonnage et publication

217. Les données sont pondérées pour la probabilité de sélection inégale et pour les cas de non-réponse unitaire. La post-stratification se fait selon la distribution par âge de la population (8 groupes), par sexe et par région (12 régions). Le fait que la post-stratification a lieu au niveau individuel signifie que les membres du même ménage peuvent recevoir des pondérations différentes.

218. Les erreurs d'échantillonnage et les effets de conception sont régulièrement estimés uniquement pour les variables clés : le taux de chômage et le ratio emploi/population. Les effets de conception sont relativement faibles; par exemple, l'effet de conception est de 1,3 pour le taux de chômage.

219. Les coefficients de variation (CV) des estimations sont calculés régulièrement. Les estimations dont le CV est inférieur à 10 % sont publiées sans aucune restriction; les estimations dont le CV est compris entre 10 et 20 % sont publiées en un seul groupe. Les CV compris entre 20 et 30 % sont publiés en deux groupes. Lorsque le CV dépasse 30 %, les résultats sont remplacés par un point (.), ce qui signifie non-zéro mais peu fiable.

220. Les résultats de l'enquête sont publiés tous les trimestres dans les Rapports statistiques rapides, l'Annuaire statistique et plusieurs autres publications slovènes. La série spéciale « Résultats de l'enquête » présente de façon détaillée les résultats de l'enquête et sa méthodologie. Les données paraissent également dans les publications d'autres organisations comme la Banque mondiale, le Fonds des Nations Unies pour l'enfance (UNICEF) et Eurostat. Des chercheurs extérieurs à l'Office de statistique analysent également les microdonnées.

Enquête sur le budget des ménages (HBS)

Historique

221. La première enquête sur la consommation des ménages a été exécutée dans les années 1960. Jusqu'en 1997, cette enquête a été effectuée selon la méthodologie relativement avancée et novatrice conçue par l'Office fédéral de statistique de la Yougoslavie. L'échantillon était un échantillon en grappes à deux phases avec stratification des UPE lors de la première phase. Les unités primaires d'échantillonnage étaient les zones d'énumération (ZE), sélectionnées selon une probabilité proportionnelle à

la taille (PPT). Lors de la seconde phase, les individus étaient sélectionnés à partir du registre central de population; leur sélection déterminait également celle du ménage. Dans chaque UPE, cinq ménages étaient interrogés. Jusqu'en 1993, on a utilisé la procédure de substitution pour produire cinq unités répondantes; toutefois, à partir de 1994, la « prise » par grappe a été portée de six à huit personnes au sein de chaque UPE, caractéristique qui nécessitait une correction supplémentaire avec pondérations. Deux types différents d'enquêtes HBS avaient lieu régulièrement : une enquête trimestrielle et une autre, d'une durée d'un an, effectuée à intervalles de cinq ans. La dernière enquête d'un an portait sur 3 270 ménages et la dernière enquête trimestrielle, sur 1 000 ménages. Pour l'enquête annuelle, les entrevues avaient lieu à la fin de l'année et portaient sur l'année tout entière, tandis que pour l'enquête trimestrielle, les ménages sélectionnés étaient interrogés quatre fois dans l'année.

Restructuration de l'échantillon de l'enquête HBS

222. La principale raison de cette restructuration était les nouvelles directives d'Eurostat (Eurostat, 1997).

223. On utilise le registre de population pour sélectionner les répondants. La sélection de ces personnes détermine également celle des ménages. Les pondérations sont utilisées pour tenir compte des probabilités de sélection inégales des personnes et des ménages. Les institutions ne sont pas incluses. L'échantillon annuel comprend 1 200 ménages répondants. Comme un échantillon est trop petit pour permettre l'application du modèle « Nordic », on combine les données des échantillons de trois années consécutives et on les applique à l'année intermédiaire. On obtient ainsi un échantillon de 3 600 ménages.

224. On applique l'allocation de façon proportionnée aux 47 strates. En raison de la taille relativement petite de l'échantillon et du grand nombre de strates, on n'effectue qu'une stratification implicite. Dans les petits peuplements (de moins de 1 000 habitants), les zones d'énumération servent d'UPE et sont sélectionnées selon une probabilité proportionnelle à leur taille (PPT). Quatre ménages répondants sont sélectionnés dans chaque UPE. Dans les villes moyennes et grandes, on applique la méthode de l'échantillonnage aléatoire simple. De ce fait, les effets de conception sont relativement faibles, d'environ 1,2 pour les variables clés. Les unités sont sélectionnées séparément pour chaque trimestre et réparties entre 12 des semaines du trimestre correspondant. La treizième semaine est consacrée à ce qui reste de travail avec les non-répondants.

Exécution

225. Des lettres sont adressées une semaine avant la première visite, accompagnées d'une incitation : une calculatrice de poche. Comme il s'agit d'une enquête continue, elle peut être effectuée avec un nombre plus restreint d'enquêteurs (20, par exemple).

226. Les enquêteurs enregistrent tous les contacts/toutes les tentatives de contact avec les ménages sur un formulaire spécial. La situation concernant le logement, le ménage et la personne de référence de chaque unité est donc très claire, ainsi que le nombre de tentatives de contact, le nombre de journaux remplis et les raisons éventuelles de non-réponse.

227. Les données sont recueillies au moyen d'un questionnaire rempli par l'enquêteur et des journaux tenus par les membres du ménage. Presque toutes les entrevues se font à l'aide d'un ordinateur (CAPI).

228. Les ménages tiennent le journal pendant 14 jours. Pendant cette période, ils inscrivent régulièrement leurs dépenses de la journée. On considère que les ménages répondent s'ils remplissent au moins le questionnaire de l'entrevue de base, car ce questionnaire doit fournir les deux tiers des données à recueillir auprès des ménages. On obtient des taux de réponse relativement élevés et stables (de l'ordre de 81 %) aux questionnaires de l'entrevue. En revanche, le taux de réponses complètes, journaux compris, est plus faible, d'environ 70 %.

Erreurs d'échantillonnage et publication

229. Si tous les journaux d'un ménage donné manquent, on entre les données en utilisant la méthode d'imputation « hot deck » des données manquantes que l'on prélève sur un ménage similaire. On utilise également la méthode d'imputation « hot deck » pour les cas de non-réponse ponctuelle. Chaque valeur manquante est remplacée par la valeur correspondante obtenue auprès d'un précédent répondant de la même classe d'imputation définie par la taille du ménage et ses caractéristiques sociodémographiques. En particulier, on remplace le revenu individuel manquant par celui d'un répondant présentant la même situation du point de vue de l'emploi et de l'instruction.

230. La méthode de calcul des pondérations de la conception et des pondérations de la post-stratification est semblable à celle de l'échantillon de l'enquête LFS. Par ailleurs, des facteurs d'expansion spécifiques sont établis pour compenser les différentes périodes de référence. Le coefficient appliqué au nouveau calcul est essentiellement le ratio de la période de référence de l'enquête (un an) à la période de référence de la variable en question. On a également besoin de pondérations spéciales lorsque l'on combine les données de trois années consécutives. Le calcul effectué pour une date donnée utilise donc les données de trois années, dont la moitié a trait à la période antérieure à cette date et l'autre moitié à la période postérieure.

231. La méthodologie de l'enquête HBS et les résultats sont décrits dans les publications citées ci-dessus pour l'enquête LFS.

Conclusions

232. Avant l'indépendance, en 1991, les enquêtes par sondage sur les ménages n'étaient pas un moyen courant de collecte de données en Slovénie. Toutefois, à la différence d'autres pays en transition, la Slovénie a effectué régulièrement des enquêtes HBS à partir du milieu des années 1960, et a commencé la série des enquêtes LFS annuelles à la fin des années 1980.

233. Après l'indépendance, l'Office de statistique de la Slovénie a subi une transition efficace et en douceur. L'Office de statistique procède à présent régulièrement aux séries standard d'enquêtes sur les ménages. Les enquêtes socio-économiques de base (LFS et HBS) sont presque entièrement harmonisées avec les normes d'Eurostat (Office de statistique de la République de Slovénie, 2001). Diverses autres enquêtes sur les ménages ont également été effectuées : l'enquête sur la consommation d'énergie des ménages [*Household Energy Consumption Survey* (HECS)], 1997; l'enquête sur l'emploi du temps [*Time-use Survey* (TUS)], 2000/2001; l'enquête mensuelle sur le comportement des consommateurs [*Monthly Consumer Attitude Survey* (CAS)]; l'enquête trimestrielle sur les déplacements de la population nationale [*Quarterly Survey on Travels of the Domestic Population* (QSTDP)]; et l'enquête annuelle sur le crime et la victimisation [*Annual Crime and Victimization Survey* (2000, 2001)].

234. Il reste place pour de nouvelles améliorations du système d'enquêtes sur les ménages. La Slovénie a un système statistique riche et précis fondé sur l'enregis-

trement (données fiscales, base de données sur les employés, bases de données sur les assurances, etc.) qui peut être lié efficacement aux systèmes géographiques et aux données du recensement. Cela offre d'autres avantages pour la conception d'échantillons optimaux ainsi que pour les estimations.

REMERCIEMENTS

Les Présidents des Offices centraux de statistique des pays suivants ont accepté de fournir des informations sur leurs enquêtes par sondage sur les ménages : le Bélarus, la Bulgarie, la Croatie, l'Estonie, la Fédération de Russie, la Hongrie, la Lettonie, la Lituanie, la Pologne, la République tchèque, la Roumanie, la Slovaquie, la Slovénie et l'Ukraine. L'auteur tient à exprimer sa gratitude à de nombreux collègues de ces pays pour les données qu'ils lui ont fournies et pour leurs précieux commentaires. Il tient également à remercier les auteurs de leurs contributions à la section B qui a pour objet de servir de complément aux informations fournies dans la section A.

L'auteur remercie également certains réviseurs anonymes ainsi que les participants à la Réunion du Groupe d'experts sur l'analyse des caractéristiques de fonctionnement des enquêtes dans les pays en développement et les pays en transition, tenue à New York du 8 au 11 octobre 2002, de leurs commentaires et suggestions fort utiles.

RÉFÉRENCES

- Bracha, Cz. (1996). Schemat losowania próby do Mikrospisu 1995 (Sample design for the 1995 microcensus). *Wiadomości Statystyczne*, n° 3, pp. 12-18.
- Deville, J.-C., et C. E. Särndal (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association*, vol. 87, pp. 376-382.
- Éltető, Ö. (2000). Enlargement of the sample of the Hungarian LFS to get reliable small area estimates for labour market indicators. *Statistics in Transition*, vol. 4, n° 4, pp. 549-563.
- _____, et L. Mihalyffy (2002). Household surveys in Hungary, *Statistics in Transition*, vol. 5, n° 4, pp. 521-540.
- Eurostat (1995). *The Future of European Social Statistics: Guidelines and Strategies*. Luxembourg.
- _____(1996). *The Future of European Social Statistics: Use of Administrative Registers and Dissemination Strategies*. Luxembourg.
- _____(1997). *Family Budget Surveys in the EC: Methodology and Recommendations for Harmonisation*, Population and Social Conditions 3. Methods E. Luxembourg.
- _____(1998a). *Labour Force Survey: Methods and Definitions*. Luxembourg.
- _____(1998b). *Labour Force Survey in Central and Eastern European Countries, Methods and Definitions (Provisional)*. Luxembourg.
- _____(2001). Meeting of the Working Party Statistics on Income and Living Conditions (EU-SILC), 10 et 11 décembre 2001. Luxembourg. Document de travail.
- Garner, T. I., et al. (1993). Household surveys of economic status in Eastern Europe: an evaluation. Dans *Economic Statistics for Economies in Transition: Eastern Europe in the 1990s*. Washington, D.C., États-Unis, Bureau of Labour Statistics, pp. 309-353.
- Główny Urząd Statystyczny (GUS) [Central Statistical Office of Poland] [1971a). *Badania statystyczne metodą reprezentacyjną w krajach socjalistycznych (les enquêtes par sondage dans les pays socialistes)*, *Biblioteka Wiadomości Statystycznych*, tome 14 (Varsovie), p. 220.,

- _____ (1971b). Wybrane problemy metodologiczne badań reprezentacyjnych (Quelques problèmes méthodologiques que posent les enquêtes par sondage). *Biblioteka Wiadomości Statystycznych*, tome 15 (Varsovie), p. 151.
- _____ (1987). *Problemy integracji statystycznych badan gospodarstw domowych* (Problèmes d'intégration des enquêtes statistiques sur les ménages). *Biblioteka Wiadomości Statystycznych*, tome 34 (Varsovie).
- _____ (1997). *Stan dostosowania polskiej statystyki publicznej do standardów Unii Europejskiej, Harmonogram prac dostosowawczych* (État d'harmonisation des statistiques polonaises avec les normes de l'Union européenne), Varsovie, Miméographe.
- _____ (1998a). *Metodologia i organizacja mikrospisów* (Méthodologie et organisation de microrecensements), *Statystyka w Praktyce*. Varsovie.
- _____ (1998b). Budzet czasu ludności 1996 (L'Enquête de 1996 sur l'emploi du temps). Dans *Studia i Analizy Statystyczne*. Varsovie.
- _____ (1999). *Metodyka badania budżetów gospodarstw domowych* (Méthodologie de l'enquête sur le budget des ménages). dans *Zeszyty Metodyczne i Klasyfikacje*. Varsovie.
- _____ (2000a). *Aktywnosc ekonomiczna ludności Polski: I kwartał 2000* (Enquête sur la main-d'œuvre en Pologne : premier trimestre 2000), *Informacje i Opracowania Statystyczne*. Varsovie.
- _____ (2001), *Zasady wyceny kosztów prac statystycznych realizowanych przez służby statystyki publicznej w roku 2003* (Principe de l'évaluation du coût des travaux statistiques dans les statistiques officielles en 2003). Varsovie. Miméographe.
- Goskomstat (2000) *Methodologicheskie polozenia po statistike* (Principe méthodologique des statistiques). Vypusk treti (3^e ed.). Moskva, pp. 294.
- Groves, R. M. (1989). *Survey Errors and Survey Costs*. New York, John Wiley and Sons.
- _____, et M. P. Couper (1998). *Non-response in Household Interview Surveys*. New York, John Wiley and Sons.
- Kish, L., et M. R. Frankel (1974). Inference from complex samples (with discussion), *Journal of the Royal Statistical Society. Series B*, vol, 36, pp. 1-37.
- Kordos, J. (1963). *Seminarium statystyczne w Wiedniu* (Séminaire de statistiques à Vienne). *Przegląd Statystyczny*, n° 2, pp. 307-310.
- _____ (1970). *Mozliwosci szerszego stosowania metody reprezentacyjnej w badaniach statystycznych krajów-członków RWPG* (Possibilités d'une application plus large des méthodes d'échantillonnage dans les enquêtes statistiques des pays membres du Conseil d'assistance économique mutuelle). *Przegląd Statystyczny*, n° 1, pp. 33-50.
- _____ (1981). *Problemy badań gospodarstw domowych: drugie spotkanie statystyków europejskich w Genewie* (Problèmes posés par les enquêtes sur les ménages : la deuxième réunion de statisticiens européens). *Wiadomosci Statystyczn*, n° 11, pp. 36-38.
- _____ (1982). *Metoda rotacyjna w badaniach budżetów rodzinnych w Polsce* (Méthode de rotation dans les Enquêtes sur le budget des ménages en Pologne). *Wiadomosci Statystyczne*, n° 9.
- _____ (1985). *Towards an integrated system of household surveys in Poland*, *Bulletin of the International Statistical Institute* (Amsterdam). Vol. 51, livre 2, pp. 13-18.
- _____ (1988a). *Jakość danych statystycznych* (Qualité des données statistiques), *Państwowe Wydawnictwo Ekonomiczne* (Varsovie), pp. 204.
- _____ (1988b). *Enquêtes sur l'emploi du temps en Pologne*. *Statistical Journal of the United Nations ECE*, vol. 5, pp. 159-168.

- _____ (1996). Forty years of the household budget surveys in Poland, *Statistics in Transition*, vol. 2, n° 7, pp. 1119-1138.
- _____ (1998). Social Statistics in Poland and its Harmonisation with the European Union Standards, *Statistics in Transition*, vol. 3, n° 4, pp. 617-639.
- _____ (2001). Some data quality issues in statistical publications in Poland. *Statistics in Transition*, vol. 5, n° 3, pp. 475-489.
- _____, B. Lednicki et M. Zyra (2002). The household sample surveys in Poland, *Statistics in Transition*, vol. 5, n° 4, pp. 555-589.
- Krapavickaitė, D. (2002). The household sample surveys in Lithuania, *Statistics in Transition*, vol. 5, n° 4, pp. 591-603.
- _____, G. Klimavicius et A. Plikusas (1997). On some estimators in cluster sampling, *Proceedings of the XXXVIII Conference of the Lithuanian Mathematical Society*, pp. 298-303.
- Kurvits, M. A., K. Söstra et I. Traat (2002). Estonian household surveys: focus on the labour force survey, *Statistics in Transition*, vol. 5, n° 4.
- Lapins, J., et E. Vaskis (1996). The new household budget survey in Latvia, *Statistics in Transition*, vol. 2, n° 7, pp. 1085-1102.
- _____, et al. (2002). Household sample surveys in Latvia, *Statistics in Transition*, vol. 5, n° 4, pp. 617-641.
- Lednicki, B. (1982). Schemat losowania i metoda estymacji w rotacyjnym badaniu budżetów gospodarstw domowych (Conception de l'échantillon et méthode d'estimation dans l'enquête par rotation sur le budget des ménages). *Wiadomosci Statystyczne*, n° 9.
- Martini, A., A. Ivanova et S. Novosyolova (1996). The income and expenditure survey of Belarus: design and implementation, *Statistics in Transition*, vol. 2, n° 7, pp. 1063-1084.
- Mihalyffy, L. (1994). The unified system of household surveys in the decade 1992-2001, *Statistics in Transition*, vol. 1, n° 4, pp. 443-462.
- Nations Unies (1964). *Recommendations for the Preparation of Sample Survey Reports (Provisional Issue)*, Statistical Papers, Series C, No. 1, Rev. 2, ST/STAT/SER/C/1/Rev.2, New York.
- _____ (1984), *Handbook of Household Surveys (Revised Edition)*, Studies in Methods, n° 31, New York, numéro de vente : E.83.XVII.13.
- _____ (2000), *Classifications of Expenditure According to Purpose: Classification of the Functions of Government (COFOG), Classification of Individual Consumption According to Purpose (COICOP), Classification of the Purposes of Non-Profit Institutions Serving Households (COPNI), Classification of the Outlays of Producers According to Purpose (COPP)*, Statistical Papers, No. 84, numéro de vente : E.00.XVII.6.
- Postnikov, S. (1953). O metodach otbora semei rabobcich, slujaščich i kolchoznikov dla obsledovanjia ich budžeta (Sur les méthodes de sélection des familles d'ouvriers, d'employés et d'agents de coopératives pour une enquête sur leurs budgets). *Vestnik Statistiki*, n° 3, pp. 14-25.
- Särndal, C-E., B. I. Swensson et J. Wretman (1992). *Model Assisted Survey Sampling*. New York, Springer-Verlag.
- Šniukstiene, Z., G. Vanagaite et G. Binkauskienė (1996). Household Budget Survey in Lithuania. *Statistics in Transition*, vol. 2, n° 7, pp. 1103-1117.
- Office de statistique de l'Estonie (1999). *Estonian Labour Force Survey 1998: Methodological Report*, Tallinn.
- Office de statistique de la République de Slovénie (2001). *Slovenian Statistical System: A Global Assessment*, 2001. A Phare project, Ljubljana.

- Szablowski, J., J. Wesolowski et R. Wieczorkowski (1996). Indeks zgodnosci jako miara jakosci danych: na podstawie wyników spisu kontrolnego do Mikrospisu 1995 (Indice d'adaptation comme mesure de la qualité des données : sur la base de l'enquête post-énumération du microrecensement, 1995). *Wiadomosci Statystyczne*, n° 4, pp. 43-49.
- Szarkowski, A., et J. Witkowski (1994). The Polish Labour Force Survey, *Statistics in Transition*, vol. 1, n° 4, pp. 467-483.
- Traat, I. (1999). *Redesign of the Household Budget Survey: Final Report of the Sampling Group*, Tallinn, Office de statistique de l'Estonie.
- _____, A. Kukk et K. Sõstra (2000). Sampling and estimation methods in the Estonian household budget survey. *Statistics in Transition*, vol. 4, n° 6, pp. 1029-1046.
- Union européenne (1998). N° 577/98 du Conseil en date du 9 mars 1998 sur l'organisation d'une enquête par sondage sur la main-d'œuvre dans la Communauté, *Journal officiel des Communautés européennes*, 14/3/98, pp. I.77/3-I.77/7.
- Vehovar, V. (1997). The Labour Force Survey in Slovenia, *Statistics in Transition*, vol. 3, n° 1 (juin), pp. 191-199.
- _____(1999). Field substitution and unit non-response, *Journal of Official Statistics*, vol. 15, n° 2, pp. 335-350.
- _____, et M. Zaletel (1995). Non-response trends in Slovenia, *Statistics in Transition*, vol. 2, n° 5, pp. 775-788.
- Verma, V. (1995). Technical Report on the Turkey Labour Force Survey. Project TUR/86/081, Genève, Organisation internationale du Travail.
- Wolter, K. M. (1985). *Introduction to Variance Estimation*, New York, Springer-Verlag.

