

DÉPARTEMENT DES AFFAIRES ÉCONOMIQUES ET SOCIALES
BUREAU DE STATISTIQUE DES NATIONS UNIES

ÉTUDES STATISTIQUES

Série **F** N° **19**

**ORGANISATION ET CONDUITE
DES ENQUÊTES
SUR LA DISTRIBUTION**



NATIONS UNIES
New York, 1977

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.

ST/ESA/STAT/SER.F/19

PUBLICATION DES NATIONS UNIES

Numéro de vente : F.77.XVII.3

Prix : \$ E.-U. 10,00
(ou l'équivalent en monnaie du pays)

TABLE DES MATIERES

	<u>Page</u>
INTRODUCTION	1
Observations générales	1
Organisation du manuel	2
 <u>Chapitres</u>	
I. DEMARCHES PRELIMINAIRES ET PRINCIPALES CONSIDERATIONS	3
Etude des programmes existants	3
Etude des recommandations internationales	4
Principales considérations	4
II. STATISTIQUES DE LA DISTRIBUTION : BESOINS ET USAGES	8
Observations générales	8
Besoins et usages officiels	8
Besoins et usages privés	9
Autres besoins et usages	9
III. PLANIFICATION ET ORGANISATION GENERALES	11
Préparation d'un exposé justificatif	11
Fondement législatif	11
Détermination de la structure organique	11
Constitution du personnel de planification	12
Organisation de groupes consultatifs	13
Décisions fondamentales	14
Méthode de réunion des données	14
Méthode de dépouillement des données	15
Besoins en matériel et en personnel	15
Répartition des responsabilités	15
Coordination avec d'autres enquêtes	16
Préparation d'un essai préalable	17
Préparation des publications	18
Différences de planification entre programmes de recensement et programmes d'enquête	19
IV. GESTION ET BUDGET	20
Utilisation des manuels	20
Emploi d'organigrammes	20
Calendriers	21
Calendrier d'affectation du personnel et du matériel ...	21
Rapports d'avancement	22
Institution d'un système de contrôle qualitatif	23
Maximisation de l'emploi des ressources	23
Différences entre recensements et enquêtes courantes ...	25

TABLE DES MATIERES (suite)

	<u>Page</u>
<u>Chapitres (suite)</u>	
V. PORTEE, CHAMP ET CLASSIFICATION	26
Introduction	26
Portée	26
Champ	28
Classification	31
VI. QUESTIONNAIRES ET AUTRES FORMULAIRES D'ENQUETE	32
Renseignements nécessaires au contrôle, à la correspondance et aux interpolations	32
Questions en vue de la publication et du collationnement	32
Distinction entre réponses omises et réponses "néant" ..	33
Rédaction de la question sur le genre d'activité	34
Détection des déclarations communes	34
Codage du formulaire	35
Importance du précodage	35
Rédaction des instructions	35
Problèmes résultant de la diversité des activités	36
Conception des formulaires	37
Succession et libellé des questions	38
VII. REUNION DES DONNEES	40
Observations générales	40
Méthodes de dénombrement	41
Dénombrement sur place	41
Constitution et utilisation d'un répertoire	45
Collecte et contrôle des déclarations	50
Recours au sondage	53
VIII. TRAITEMENT DES DONNEES	56
Observations générales	56
Simples opérations de bureau	56
Préparation des questionnaires en vue de les tabuler ...	58
Collationnement et codage par ordinateur	61
Données provenant de dossiers administratifs	63
Collationnement des enquêtes itératives par sondage	64

TABLE DES MATIERES (suite)

	<u>Page</u>
Chapitres (suite)	
IX. MISE EN TABLEAUX, REVISION ET CORRECTION	65
Observations générales	65
Machines à additionner	65
Traitement électromécanique	65
Traitement informatique	67
Révision et correction des tableaux	68
X. PUBLICATION	72
Révision finale	72
Analyse finale des risques d'indiscrétions	72
Préparation et reproduction des exemplaires	72
Publication de rapports préliminaires	73
Renseignements à publier au sujet des enquêtes	73
Contenu des tableaux	75
Enquêtes ayant lieu plusieurs fois par an	79
ANNEXES	
I. TECHNIQUES DE SONDAGE DANS LES ENQUETES SUR LA DISTRIBUTION ET LES SERVICES	80
II. BIBLIOGRAPHIE	106

INTRODUCTION

Observations générales

1. La Commission de statistique des Nations Unies, lors de sa dix-huitième session tenue en 1974, a notamment adopté des recommandations sur les statistiques de la distribution et des services, faisant sienne une proposition du Secrétaire général et recommandant "l'établissement d'un manuel sur l'organisation et l'exécution des enquêtes qui serait publié en tant que supplément méthodologique aux recommandations internationales" 1/.
2. Préparée pour donner suite à cette recommandation, la présente publication est destinée à servir de guide pratique à l'aménagement et à la conduite des recensements et enquêtes sur la distribution et les services; elle s'attache particulièrement aux problèmes qui se posent aux pays où débute le développement statistique.
3. Le cadre conceptuel des enquêtes, dont le présent manuel expose l'exécution dans ses aspects pratiques, se trouve tracé dans les recommandations internationales adoptées par la Commission de statistique 2/.
4. Vu les analogies entre les enquêtes sur l'industrie et celles sur la distribution, le présent manuel contient un exposé de certains procédés et méthodes qui s'appliquent aussi aux enquêtes industrielles. Toutefois, comme actuellement la plupart des pays enquêtent séparément sur l'industrie et sur la distribution, on estime que des manuels différents, par leur caractère plus spécialisé, serviront plus aisément de guides aux pays qui envisagent de mener de telles enquêtes.
5. Aux fins du présent manuel, on entend par recensement une enquête générale réalisée à longs intervalles, d'ordinaire cinq ou dix ans. Les enquêtes courantes sont au contraire menées à intervalles plus rapprochés : elles sont mensuelles, trimestrielles ou annuelles. Recensements et enquêtes embrassent les activités du commerce de gros et de détail, des restaurants et hôtels et de certains services, définies dans la Classification internationale type, par industrie, de toutes les branches d'activité économique (CITI) 3/, comme formant la branche 6 et certaines parties des catégories 83, 94 et 95. Les unités pour lesquelles ces statistiques sont établies sont les établissements commerciaux (magasins,

1/ Conseil économique et social, Documents officiels : Cinquante-huitième session, supplément No 2 (E/5603), par. 97.

2/ Recommandations internationales pour les statistiques de la distribution et des services, Etudes statistiques, série M, No 57 (publication des Nations Unies, numéro de vente E.75.XVII.9).

3/ Classification internationale type, par industrie, de toutes les branches d'activité économique, Etudes statistiques, série M, No 4, Rev.2 (publication des Nations Unies, numéro de vente E.68.XVII.8).

boutiques, bureaux et installations analogues) situés dans les limites territoriales du pays considéré. L'entreprise, tout en étant dans certains cas l'unité déclarante, n'est donc pas l'unité statistique de base.

Organisation du manuel

6. Le présent manuel se propose d'exposer la marche à suivre pour aménager et conduire une enquête sur la distribution. Sa matière a été répartie, chaque fois que possible, dans l'ordre où doivent s'envisager les opérations. A maints égards, les méthodes servant aux recensements espacés et aux enquêtes courantes sont parfaitement analogues. Toutefois, quand elles diffèrent, le texte l'indique.

7. Le chapitre I décrit les démarches préliminaires et les principales considérations en vue de l'organisation des enquêtes; le chapitre II expose la nécessité et les usages des données sur la distribution; les chapitres III et IV sont consacrés à l'étude de l'aménagement et de l'organisation générale, de l'exécution et du budget. Les chapitres V et VI examinent les questions relatives à la portée, au champ, à la classification et à la conception des formulaires et questionnaires. Le chapitre VII envisage les méthodes et problèmes de la réunion des données. Les chapitres VIII et IX décrivent les démarches et problèmes qu'entraînent le traitement des données, la présentation en tableaux, la révision et la correction. Le manuel s'achève avec le chapitre X, qui traite de la publication des résultats. L'annexe I contient une étude détaillée des techniques pratiques de sondage applicables aux enquêtes sur la distribution.

8. On ne doit voir dans le manuel qu'un ensemble de suggestions pratiques sur les principaux aspects de l'organisation et de l'exécution d'enquêtes sur la distribution. Son cadre restreint n'a pas permis d'exposer chaque sujet avec tout le détail voulu et les recommandations qui y figurent ne sont qu'un guide applicable à certains aspects de l'exécution d'une enquête. Tout pays qui envisage d'en entreprendre devra approfondir ces recommandations pour les adapter à ses conditions particulières.

9. Bien que les recommandations internationales visent à traduire les pratiques courantes dans la majorité des pays, ceux qui s'intéressent à un exposé plus détaillé de certains sujets ou au cas particulier de certains pays devront prendre contact avec ces pays pour s'informer des problèmes qu'ils peuvent avoir rencontrés et pour profiter de leurs solutions. On pourra aussi se renseigner sur les problèmes particuliers en écrivant au Directeur du Bureau de statistique, Nations Unies, New York, N.Y. 10017.

I. DEMARCHES PRELIMINAIRES ET PRINCIPALES CONSIDERATIONS

Etude des programmes existants

10. Tandis que les recommandations internationales proposent le genre de données à réunir dans le secteur de la distribution et des services, il incombe à chaque pays d'examiner pour son compte les données qu'il collecte déjà et la nature de ses objectifs. A cet effet, il convient d'étudier les rapports administratifs réunis par les divers organismes officiels et de s'informer soigneusement des données nécessaires aux secteurs public et privé.

11. Un examen des activités officielles peut révéler qu'on dispose déjà d'une mine de données grâce aux services du fisc, à la réglementation des établissements financiers, à l'application du droit du travail ou aux services sociaux. Ces données doivent être examinées pour déterminer si elles peuvent servir à des programmes statistiques en même temps qu'aux fonctions réglementaires ou administratives qui les ont rendu nécessaires.

12. Pour déterminer les besoins, il sera nécessaire de s'adresser à deux grandes catégories d'usagers, les officiels et les particuliers. Il est bon de noter que gouvernants et dirigeants publics décident souvent sans renseignements statistiques appropriés. Auraient-ils la possibilité de faire connaître leurs besoins en cette matière, leurs demandes seraient énormes. Le problème consisterait alors à limiter le programme aux possibilités.

13. Un moyen de définir les besoins des particuliers serait de voir les programmes qu'entreprennent les organismes privés. On constaterait d'ordinaire l'impossibilité de se fier statistiquement à ces programmes, qui ne parviennent pas à s'assurer des réponses suffisamment nombreuses et représentatives, mais qui sont réalisés faute de pouvoir tenter toute autre démarche. Que les organismes considérés soient néanmoins poussés à cet effort montre assez l'utilité des données.

14. Dans l'examen des besoins, il convient de ne pas oublier que deux grandes catégories de programmes, d'ailleurs largement coordonnés, doivent être considérées. La première consiste en un recensement général qui cherche à élaborer des statistiques en suffisamment grand détail pour servir à comprendre la structure du secteur tout entier. Un tel recensement a d'ordinaire lieu à longs intervalles, peut-être tous les cinq ans ou tous les dix ans; sa préparation et son exploitation exigent beaucoup de temps. Le second type de programme a lieu à intervalles plus courts : il est mensuel, trimestriel ou annuel; sa matière est bien plus réduite et il peut s'opérer par sondage. Les échantillons utilisés ne doivent pas obligatoirement provenir de dénombremments censitaires complets; mais s'ils en sont tirés, ils seront plus sûrs et plus utiles. Leur origine permet alors de lier les deux séries de données et sert à contrôler le déroulement des sondages intercensitaires.

Etude des recommandations internationales

15. Comme on l'a dit plus haut, les recommandations internationales adoptées par la Commission de statistique et publiées depuis par les Nations Unies ^{4/} sont à la disposition des pays intéressés pour les aider à développer leurs statistiques de la distribution. Ces recommandations traitent des sujets suivants : portée et champ des enquêtes et unités à utiliser; caractéristiques de l'unité statistique et classification à établir; catégories de données à rassembler et statistiques à publier; et définitions des catégories de données. Le programme ainsi esquissé comprend des recommandations concernant les enquêtes faites à grands intervalles (recensements), les enquêtes annuelles et celles qui ont lieu plusieurs fois par an. Pour chaque catégorie, des recommandations concernent les rubriques à inclure. Sont également traitées les principales difficultés, notamment la définition de l'unité statistique et de l'unité déclarante, et celle des caractéristiques de l'établissement, telles que type d'activité, nature de l'établissement, type d'organisation économique, type d'organisation juridique et régime de propriété, emplacement et dimension. On n'escompte pas que ces recommandations soient jugées acceptables dans tous leurs détails par chaque pays. Des modifications appropriées doivent leur être apportées selon la situation particulière de chacun. Toutefois, on devra s'efforcer de faire entrer le programme national dans le cadre des recommandations internationales, de façon à pouvoir se livrer à d'utiles comparaisons de pays à pays et à pouvoir développer davantage les statistiques mondiales fondées sur des notions types. A mesure que l'économie mondiale devient de plus en plus coordonnée et interdépendante, pareille normalisation peut conduire à une meilleure compréhension du jeu mutuel des forces économiques et à une solution plus intelligente des problèmes économiques.

Principales considérations

16. Etant donné la grande complexité tant des recensements espacés que des enquêtes fréquentes, il importe que tous les aspects de la tâche à accomplir soient compris et préparés. Il est nécessaire d'examiner soigneusement toutes les opérations et leurs effets réciproques pour s'assurer que le résultat final sera acceptable. Il en va particulièrement ainsi des recensements espacés, pour lesquels il faudra attendre cinq ou dix ans avant de pouvoir modifier les instructions. Pour les enquêtes plus fréquentes, les corrections doivent être apportées plus souvent et l'erreur éventuelle sera moins désastreuse. Il importe d'élaborer un programme bien conçu et spécialement adapté aux besoins du pays considéré. Des systèmes statistiques qui fonctionnent bien dans un pays peuvent être d'usage tout à fait inapproprié dans un autre.

^{4/} Recommandations internationales pour les statistiques de la distribution et des services, Etudes statistiques, série M, No 57 (publication des Nations Unies, numéro de vente &.75.XVII.9).

17. Un budget pertinent et réaliste dépend d'une préparation soigneuse et d'une juste analyse des coûts. Souvent, au responsable des aspects techniques du programme incombe aussi la charge d'établir le budget, d'où résulte une sous-estimation des dépenses. Il importe qu'un analyste budgétaire compétent examine toutes les opérations pour s'assurer que le budget est approprié et suffisant. Des réserves pour imprévus doivent être attribuées en cas de besoin. Tout le déroulement de l'opération doit être soigneusement analysé pour qu'on soit sûr de n'avoir rien oublié d'important. Par suite d'une préparation défectueuse des plans et des budgets, dans certains pays la seule réunion des données a largement absorbé les crédits qui ont manqué ainsi pour le dépouillement des questionnaires et la publication des résultats.

18. Un défaut courant, surtout dans les recensements espacés, provient d'une surcharge du programme et d'une complication du questionnaire portées au point que toute l'opération se trouve bloquée par sa complexité même. S'il importe certes de s'informer de l'essentiel, cette information risque d'être perdue, soit que les recensés se découragent devant la quantité des questions posées ou ne disposent pas aisément de certaines réponses, soit qu'il devienne impossible de dépouiller des questionnaires si poussés. Il importe particulièrement, en lançant un programme nouveau, de simplifier le plus possible la matière du questionnaire, pour assurer la réussite du recensement.

19. Pour qu'une enquête réussisse, il est essentiel de s'assurer la coopération des enquêtés. Une façon d'y parvenir est d'en faire une obligation légale; et l'application de sanctions pour abstention s'impose probablement dans le cas des recensements généraux. Les sanctions apparaissent moins justifiées pour les enquêtes courantes par sondage, puisqu'il y est demandé aux sondés d'assumer une tâche qui n'est pas requise de tous. De toute façon, une participation spontanée assurera mieux qu'une participation forcée une information correcte et ponctuelle. Il incombe donc au responsable de l'enquête de s'entendre avec les entreprises commerciales et les représentants des grandes sociétés pour créer un climat approprié à la conduite de l'enquête. Parfois, un tel climat dépendra de l'inclusion d'une rubrique particulière demandée par une branche déterminée et parfois de l'exclusion d'une rubrique inopportune. Ce climat exige aussi que les données requises puissent facilement se tirer des registres comptables des enquêtés. Les entreprises se plaignent assez souvent que les données requises ont déjà été demandées et communiquées à une autre administration. Elles peuvent avoir raison, mais il est souvent nécessaire de redemander les mêmes renseignements pour une question de temps ou vu la difficulté de les obtenir de cette autre administration.

20. Comme on l'a dit, il est possible d'utiliser des organes de tutelle aux fins de l'enquête. D'abord, ils disposent généralement d'une liste d'adresses, ou peuvent la tirer de leurs registres. Pareille liste contient fréquemment d'autres indications que la classification industrielle et parfois quelques données, comme le volume des ventes. De tels renseignements sont très utiles pour élaborer un répertoire qui servira au recensement ou à préparer les échantillons destinés aux enquêtes courantes. Ensuite, si l'organe de tutelle réunit déjà des données, on peut lui demander de modifier ses demandes pour les adapter aux besoins de l'enquête.

21. L'emploi des registres d'autres administrations exigera fréquemment d'adroites négociations, car ces administrations ne se préoccupent guère de statistique. Il exigera aussi une préparation soigneuse pour obtenir que les registres soient disponibles en temps opportun et dans la forme requise par l'enquête. Un problème peut se poser si les données provenant d'un organe de tutelle ont été obtenues par lui à titre confidentiel. Etant donné que les enquêtes réuniront aussi des données en vertu d'une loi qui garantit le même secret, l'emploi des données provenant d'un organe de tutelle peut se justifier pour autant que le personnel du service statistique soit astreint au secret professionnel et que les publications ne contiennent rien qui dévoile les activités d'une entreprise quelconque. L'inverse toutefois n'est pas vrai; en effet, les renseignements sur telle ou telle entreprise réunis au cours d'une enquête statistique ne peuvent être communiqués à un organe de tutelle, car ils ont été obtenus contre la promesse qu'ils ne serviraient pas à des fins réglementaires ou fiscales.

22. Nombre de rubriques figurant dans les enquêtes sur la distribution et les services sont analogues à celles comprises dans d'autres enquêtes menées soit par le service statistique, soit par d'autres organes officiels. Il importe de veiller dans toute la mesure du possible à définir de la même manière ces rubriques communes, de façon qu'on puisse comparer ou additionner les données provenant de diverses enquêtes. Les modifications qui peuvent s'imposer à cause des conditions particulières d'une branche donnée doivent être conçues en vue de rendre comparables les données qui en résultent. Cette considération importe particulièrement pour les données qui sont réunies à la fois dans un recensement général et dans une enquête courante. Par exemple, vu la nécessité de réunir et de publier rapidement les données d'une enquête mensuelle, l'enquêteur peut éprouver quelque difficulté à calculer les données relatives à ses ventes pour des rubriques comme la taxe à la valeur ajoutée, les reprises, les remboursements, etc. Il peut en résulter un manque de comparabilité entre recensements et enquêtes courantes.

23. Dans la plupart des pays, existe déjà d'ordinaire un service de réunion des données. On devra alors décider dans quelle mesure on peut ou doit intégrer les diverses enquêtes, car elles comprennent d'ordinaire nombre d'éléments communs. S'il est possible d'intégrer maintes tâches d'exploitation, celles de planification, surtout s'il s'agit de la matière et de la connaissance de la branche, devraient être assurées à part. Même pour les tâches d'exploitation, on doit veiller à ne pas oublier qu'il importe de connaître les branches considérées pour estimer la nature des problèmes et la façon de les résoudre.

24. Bien des enquêtes se relèvent infructueuses malgré les grands efforts consentis avant leur réalisation, car c'est à leur conduite effective que s'attache leur succès. Dans tous les cas, avant de lancer une enquête, il convient pour le moins de rendre visite aux entreprises pour examiner avec elles le contenu du questionnaire. Toutefois, il vaut beaucoup mieux procéder à un essai préalable en grandeur réelle pour tester tous les aspects de l'enquête, de sa préparation à la publication des résultats. C'est ce qu'expose plus en détail le paragraphe 65.

25. Les recommandations internationales publiées par les Nations Unies^{5/} contiennent un bref exposé de la nécessité des statistiques de la distribution. Le chapitre suivant s'efforce de présenter les principaux usagers des renseignements demandés dans les enquêtes. Il faut espérer que la pensée de ces usagers, de même que des fins pour lesquelles ils ont besoin des renseignements, aidera ceux qui préparent une enquête non seulement à choisir les rubriques à y inclure et à décider de leur importance, mais aussi à assurer la publicité et à répondre aux questions sur l'utilité du programme.

^{5/} Recommandations internationales pour les statistiques de la distribution et des services, Etudes statistiques, série M, No 57 (publication des Nations Unies, numéro de vente E.75.XVII.9).

II. STATISTIQUES DE LA DISTRIBUTION : BESOINS ET USAGES

Observations générales

26. Les statistiques de la distribution et des services ont tendance à être réunies bien plus tard dans le cours du développement économique que celles sur l'agriculture, les industries manufacturières et les industries extractives. Il en va ainsi parce que les activités fondamentales de l'homme visent d'abord à satisfaire ses besoins essentiels, à savoir nourriture, logement et vêtement. C'est seulement quand la société se trouve en mesure d'accumuler des excédents que le commerce prend de l'importance. La spécialisation sous forme de fourniture de services intervient plus tard. De même qu'on avait trouvé utile de réunir des renseignements au sujet de l'agriculture et des manufactures pour aider à comprendre et à améliorer ces activités, il devient utile de réunir des données sur la distribution et les services quand ils deviennent aussi des branches importantes.

Besoins et usages officiels

27. Les diverses administrations publiques peuvent utiliser à bien des fins les données sur la distribution et les services. Une des principales est comme base d'estimation des tendances économiques. Pour les tendances à court terme, les estimations annuelles, ou plus fréquentes encore, de ce secteur de l'économie nationale jouent un rôle important. Il en va particulièrement ainsi pour le commerce de détail et les services personnels, car ces activités produisent la consommation finale, objet de toute l'activité économique. Dans les économies de marché, la hausse et la baisse de la consommation finale influent nécessairement sur tout le reste de l'activité économique et deviennent ainsi le principal indicateur à observer par les services officiels chargés de la politique économique. Dans la mesure où production et consommation ne correspondent pas par le moyen des échanges, cet indicateur perd de son importance. Outre la consommation finale, il aide à mesurer d'autres aspects de la distribution et des services, tels que salaires et traitements gagnés dans ce secteur, mouvements des stocks, formation brute de capital fixe.

28. Outre mesurer les tendances de l'économie, les données sur la distribution et les services peuvent servir à définir les mesures à prendre (ainsi qu'à mesurer les effets de celles déjà en vigueur) dans de nombreux secteurs différents de l'activité de l'Etat : par exemple, pour estimer le produit qu'on peut attendre de l'imposition d'une fiscalité telle que l'impôt sur les ventes, les taxes à la valeur ajoutée et l'impôt sur les salaires et pour déterminer si ce produit a été atteint. Des données sont aussi requises pour calculer les statistiques du revenu et du produit national. Elles peuvent encore contribuer à d'autres programmes, comme l'évaluation du besoin de modifier le système de transport, la mesure du besoin en main-d'oeuvre et de son utilisation, l'aménagement des programmes de développement régional, la rédaction des programmes d'action sociale tels que sécurité sociale, salaire minimum et assurance-chômage, l'évaluation des programmes antitrust, etc.

29. Malgré l'utilité pour l'Etat des données sur la distribution et les services, on ne doit jamais oublier que fournir ces informations entraîne pour l'entreprise des frais qui au total peuvent être bien plus élevés que pour le gouvernement. Demander des données complexes à ce secteur de l'économie peut lui imposer une charge particulièrement lourde, car l'entreprise ordinaire est de petite taille et peut ne pas tenir de livres appropriés. Ces faits doivent être mûrement pesés pour déterminer les données demandées par le programme.

Besoins et usages privés

30. Principaux fournisseurs de données sur la distribution et les services, les milieux économiques en sont aussi un grand utilisateur. Ils peuvent aussi bénéficier de leur application appropriée, car les décisions prises par l'Etat et qui les affectent en seront plus judicieuses. Leurs représentants peuvent favoriser ou combattre ces décisions de façon fondée s'ils connaissent les faits. S'il est vrai que seules les grandes sociétés peuvent disposer du personnel nécessaire pour analyser les données, les petites peuvent jouer un rôle efficace dans la détermination de la politique générale par les associations professionnelles qui les représentent.

31. L'un des principaux usages de ces données par les entreprises consiste à déterminer l'emplacement de leurs points de vente. Certes, il leur faut à cet effet des statistiques non seulement sur l'économie, mais aussi sur le caractère de la population et la répartition des revenus. Les analystes industriels peuvent en tirer un bon tableau de toute une branche en ce qui concerne la taille des entreprises, leur proximité des centres habités, le rapport numérique de la masse des salaires au total des ventes ainsi que des stocks aux ventes, l'implantation sur le marché, les tendances de l'industrie dans le temps, le genre de débouchés établis pour des articles déterminés, ainsi que plusieurs autres facteurs utiles, selon le genre de données réunies et le temps et les efforts consacrés à l'analyse des renseignements ainsi rassemblés. Les renseignements plus complexes ne peuvent certes se réunir lors d'enquêtes annuelles ou plus fréquentes encore, mais seulement par des recensements généraux espacés.

32. Outre les entreprises d'une branche déterminée qui les utilisent, les données peuvent se révéler utiles à d'autres dans des branches qui fournissent des services. Par exemple, les services publics et les compagnies de transport dresseront bien mieux leurs plans s'ils ont une connaissance détaillée de la structure des nombreuses unités qui interviennent dans le commerce de gros et de détail et les services. De même, toutes les catégories de fabricants d'équipement peuvent étudier le marché de leurs produits. Les fabricants d'articles qui seront vendus par des grossistes ou des détaillants peuvent aussi en tirer profit. Les établissements financiers forment une autre catégorie fortement intéressée par les statistiques.

Autres besoins et usages

33. Une utilisation à ne pas négliger est celle qui est faite dans l'enseignement. En premier lieu, les données servent dans les instituts d'administration des affaires à former les étudiants qui entreront comme cadres dans des entreprises grandes et moyennes. Par une autre utilisation,

elles servent à former les économistes, les analystes de marché et les statisticiens spécialisés dans les enquêtes. Ceux qui étudient pour devenir économistes du travail, conseillers financiers ou économistes des services publics les trouvent aussi utiles.

34. Aux organisations ouvrières, elles seront de même utiles pour leur rôle en matière de négociations de salaire. Ces organisations peuvent comparer les données sur l'effectif de leurs adhérents avec des renseignements sur la répartition et l'effectif des travailleurs dans les secteurs dont elles s'occupent, afin de déterminer les points les plus propices à leurs efforts d'organisation des travailleurs. De plus, la connaissance qu'elles s'assurent de la structure et du fonctionnement d'une branche d'activité peut les aider à présenter des demandes légitimes et à négocier des conventions efficaces.

35. Les données provenant de recensements généraux espacés peuvent trouver une utilisation fondamentale dans l'établissement des plans de sondage qui serviront à des enquêtes, notamment annuelles, ou plus fréquentes encore. Le recensement général fournit le cadre idéal d'où tirer l'échantillon le plus efficace, puisqu'il fournit les caractéristiques des unités destinées à être sondées, telles que nature, taille et emplacement des entreprises, et qu'il procède à un dénombrement complet comprenant le nom et l'adresse de chacune. Bien que les organismes privés de sondage ne puissent avoir communication des renseignements détaillés dont dispose le gouvernement, il leur reste possible d'utiliser judicieusement les données publiées, tant pour dresser le plan d'un sondage efficace que pour vérifier leurs résultats en les comparant aux totaux officiels.

36. Enfin, l'élargissement des programmes statistiques, par la diffusion des renseignements concernant la distribution et les services, donnera une dimension nouvelle au fonds général des connaissances et donnera en public cette meilleure conscience des faits de la vie économique qui permet aux particuliers d'influer de façon éclairée sur la politique officielle. Ceux qui s'occupent de former l'opinion publique par les divers moyens de communication - journaux, périodiques, télévision - auront accès aux données qui permettent une meilleure estimation des programmes publics et privés.

III. PLANIFICATION ET ORGANISATION GENERALES

Préparation d'un exposé justificatif

37. Aux fins du présent manuel, on suppose ici que quelques responsables autorisés ont décidé provisoirement qu'il serait bon d'entamer un programme de statistiques de la distribution et des services. On suppose aussi qu'existe déjà un service statistique qui s'occupe d'autres enquêtes. Pour première démarche dans les formalités d'obtention d'approbation pour le programme, il serait bon de préparer un exposé ou dossier justificatif. Une bonne partie des mesures énumérées ci-dessous peut y servir. Ce document devrait donner quelque idée générale du coût, du temps et du personnel requis. Il n'a pas lieu de détailler beaucoup, mais devrait suffire pour donner au directeur du service quelque idée de l'ordre de grandeur, afin qu'il puisse décider s'il recommande d'approuver ou de modifier le programme. Une présentation bien faite à ce stade contribuera beaucoup à faire approuver le projet.

Fondement législatif

38. Quand le service statistique a décidé de soutenir le programme, il est d'ordinaire nécessaire de voter une loi (ou de modifier la législation en vigueur) pour lui donner une base légale. Le texte de loi doit mentionner le champ (assez vaste pour comprendre tout le commerce de gros et de détail et les services), la fréquence des recensements généraux, la nature des enquêtes annuelles et des enquêtes plus fréquentes, les règles du secret régissant la communication des renseignements fournis par une entreprise et les sanctions prévues en cas d'abstention, ou de falsification dolosive des données. Ainsi se trouve tracé le cadre juridique dans lequel les diverses enquêtes pourront être aménagées et réalisées. Quand un fonctionnaire dispose déjà de pouvoirs étendus pour réunir des renseignements économiques, un décret (ou une ordonnance) constitue une base légale suffisante. Il doit comprendre de même les dispositions énumérées ci-dessus.

Détermination de la structure organique

39. Toute décision sur la structure organique dépendra de la nature de l'organisme statistique existant. Quand existe déjà tout un ensemble fonctionnalisé, avec des bureaux séparés par matière, un personnel de préparation, des services extérieurs, une division chargée des publications, etc., on peut se borner à lui adjoindre un personnel chargé d'aménager l'enquête sur la distribution et les services, qui préparera les instructions techniques destinées aux diverses divisions d'exécution, lesquelles, à leur tour, constitueront des unités distinctes pour la conduite de ces enquêtes. Quand, en revanche, les tâches d'exploitation n'ont pas été organisées séparément, il faudra que la préparation et l'exploitation se combinent en un seul groupe.

40. Une activité normalement confiée à un personnel distinct est celle qui se déroule sur le terrain et qui, selon toute probabilité, s'exécutera pour un certain nombre de programmes différents. Ce service extérieur prend une importance particulière quand le programme prévoit la réunion de données à

partir d'un sondage aréolaire dans une série d'enquêtes courantes. Si un pays ne dispose pas encore de services extérieurs, le lancement d'une telle opération devient une tâche considérable et coûteuse.

41. Pour comparer les avantages d'un système qui confie les fonctions d'exploitation et de planification à un même service et ceux d'une division du travail fondée sur la fonction à accomplir, deux facteurs importants sont la nature du personnel disponible et le volume de travail à exécuter. La fonction de planification exige un personnel rompu aux problèmes posés par le choix des matières. La fonction d'exploitation exige aussi des connaissances et compétences spécialisées. L'idéal serait que le personnel soit expert dans ces deux domaines, mais souvent ce n'est pas le cas. De plus, chaque activité absorbera probablement tout le temps et l'attention du personnel qui en est chargé.

42. Il importe extrêmement de comprendre l'effet que l'exploitation peut avoir sur la qualité des données à obtenir. Trop souvent, un personnel principalement chargé du dépouillement utilisera des expédients et apportera des changements de méthode qui, s'ils peuvent hâter le résultat et en réduire le coût, influenceront sur les données au point de les rendre inutilisables. En revanche, à réunir les deux fonctions, on risque qu'un personnel extrêmement compétent quant aux matières cause une exploitation parfaitement inefficace et un gaspillage de crédits qui entraîne l'abandon du programme. Il est essentiel de garder en tout temps les deux fonctions présentes à l'esprit et de s'assurer que ceux qui s'occupent d'élaborer les plans envisagent toujours la façon dont ils influenceront sur l'exploitation et inversement. Au personnel d'exploitation devrait être donné un cours d'initiation aux matières, auquel ferait pendant un bref cours d'exploitation donné au personnel chargé du choix des questions.

Constitution du personnel de planification

43. Les fonctions incombant au personnel de planification consistent à déterminer les matières qui seront retenues; à consulter les utilisateurs de données, les associations professionnelles et les enquêtés; à concevoir et à essayer les questionnaires; à participer à l'aménagement de la réunion des données et des méthodes de dépouillement; à préparer des instructions destinées aux enquêtés; à préparer les instructions techniques applicables au dépouillement; à résoudre les problèmes posés par la réunion des données; à préparer les publications et à rédiger les textes; et à revoir, quant au fond, les tableaux des données. Le personnel chargé de ces fonctions doit manifestement être fort capable. Les compétences essentielles requises sont la connaissance de la statistique et de l'économie, avec quelque aptitude aux mathématiques.

44. Un programme intégral comprenant à la fois des recensements espacés et des enquêtes plus fréquentes exigera de préférence la mise sur pied de ceux groupes de planification, un par programme. En effet, un programme d'enquêtes fréquentes et suivies, correspondant par exemple à des séries mensuelles, absorbe toute l'attention du personnel qui en est chargé; lui confier de surcroît les tâches relatives au recensement général nuirait aux deux programmes.

45. Au personnel chargé d'aménager des phases spécifiques du recensement doit ensuite incomber la réalisation. On contribue ainsi à la continuité dans l'aménagement et l'exécution des opérations prévues. De plus, la formation et l'expérience acquises lors de la phase de planification servent à plein lors des opérations ultérieures.

Organisation de groupes consultatifs

46. Dès le début, le personnel de planification devra organiser deux grands groupes consultatifs d'utilisateurs. Il profitera ainsi du jugement et des connaissances de l'ensemble des fonctionnaires intéressés et du public, qui l'aideront beaucoup à déterminer les objectifs du recensement. Ces groupes aideront aussi à lancer le recensement en contribuant à l'obtention du cadre légal, des crédits et de la coopération publique.

47. Le premier groupe à former ainsi est un groupe d'officiels. Il sera composé de fonctionnaires expérimentés, qui utilisent déjà ou utiliseront les données considérées. Ces fonctionnaires doivent occuper des postes élevés, de manière à pouvoir parler de façon autorisée pour les services qu'ils représentent. Il doit s'agir aussi de techniciens qui comprennent la nature des programmes statistiques et leurs limites. Ils peuvent venir, par exemple, des ministères du travail, de la sécurité sociale, du commerce et d'autres administrations chargées de réglementer ou de favoriser le commerce et les services, ou encore d'assurer la planification et l'analyse économiques.

48. Les usagers privés forment la seconde catégorie pour laquelle il y a lieu de former un groupe consultatif. Y figurent des représentants des groupes d'analyse des marchés (pratique et théorie), des représentants des moyens de transmission (journaux, radio, télévision), des organisations d'études statistiques, des économistes et des organisations de consommateurs. Là encore, on doit s'efforcer de choisir des membres qui aient quelque connaissance des problèmes posés par les enquêtes et puissent contribuer à leur aménagement, de préférence à des personnes qui ne pourraient s'acquitter de ces fonctions. S'il se révèle que le comité consultatif se trouve gêné par sa taille, on pourra devoir le diviser en deux ou plusieurs sous-groupes.

49. Deux types d'organismes ont été délibérément exclus des groupes consultatifs, non qu'on n'ait besoin de leurs avis, mais parce que leur consultation présente un caractère différent. Le premier comprend les associations professionnelles, qu'il y a lieu de consulter plus à fond et en se rapportant expressément aux besoins et problèmes relatifs à une profession déterminée. Il est donc recommandé de contacter séparément chaque association professionnelle représentant une branche comprise dans l'enquête, pour examiner avec elle l'effet de l'enquête sur cette branche. Le second groupe exclu comprend les comptables. Il y a lieu de consulter ses représentants au sujet des données disponibles et du temps nécessaire à leur déclaration. Cette consultation doit aussi se faire à part, les comptables ne devant pas siéger dans les comités consultatifs.

50. Quand le personnel chargé de la préparation est parvenu à une conclusion préliminaire quant aux matières à inclure dans l'enquête, il serait

bon de mettre sur pied un système permettant de tenir un registre officiel de toutes les filières suivies pour parvenir à déterminer la matière de l'enquête. Ce registre devrait comprendre les données du contact, la personne contactée, la méthode utilisée (courrier, entretien direct, téléphone), les observations du commentateur et la suite donnée par le représentant du personnel de préparation. Un tel registre se révélera très utile à l'avenir.

Décisions fondamentales

51. L'enquête étant approuvée et le personnel chargé de la préparer étant réuni, ce personnel doit très tôt régler au moins provisoirement plusieurs questions. Les paragraphes suivants les examinent brièvement.

Portée

52. Bien que la décision générale au sujet des branches à inclure dans l'enquête ait déjà été prise, il n'en reste pas moins nécessaire que le personnel chargé de la préparation examine cette décision à propos de chaque branche prévue et envisage les modifications qui peuvent s'imposer. Il pourra apparaître que, pour un faible surcroît de dépenses, l'enquête peut être développée de façon à traiter une branche de manière plus détaillée ou à en inclure une de plus. Ou il peut se révéler nécessaire d'éliminer une branche donnée, vu l'impossibilité de la dénombrer.

Champ couvert

53. L'une des décisions principales à prendre d'abord, qui exerce une grande influence sur le budget, concerne le champ couvert. Il importe particulièrement de décider si l'enquête se fera sans exceptions ou si, au contraire, elle omettra par exemple les exploitations à domicile, celles qui n'occupent aucun salarié, celles au-dessous d'une taille ou d'un chiffre d'affaires donnés, etc.

Crédits disponibles

54. Le personnel chargé de la planification doit savoir dans quelle mesure des crédits sont assurés à l'enquête et l'ampleur à envisager. Le service de statistique doit être en état d'en donner une indication assez précise, de façon que les objectifs fixés au programme correspondent à une estimation raisonnable des crédits qui, seront mis à disposition.

Méthode de réunion des données

55. Quand ont été prises les décisions sur la portée et le champ et qu'on peut raisonnablement compter disposer de crédits appropriés, il y a lieu d'arrêter la méthode qui servira à réunir les données. A cet effet, il faut répondre à des questions telles que : Dans quelle mesure obtiendra-t-on des données d'autres administrations publiques ? Dans quelle mesure recourra-t-on au sondage ? Quelle partie de l'univers statistique considéré pourra être traitée par correspondance ? Dans quelle mesure les données seront-elles réunies par dénombrement ? Comment déterminer la résidence

des abstentionnistes ? Comment obtenir les adresses pour l'expédition par la poste ? Les représentants du personnel tant de planification que d'exploitation devront travailler conjointement à ces questions.

Méthode de dépouillement des données

56. Avec les progrès constants réalisés dans la conception du matériel de traitement, le nombre d'options offertes pour exploiter et présenter les données continue à croître. Ici, un mot d'avertissement s'impose. Trop souvent, on décide d'utiliser un matériel électronique perfectionné, qui est en fait non seulement superflu mais encore inopportun. Même les pays les plus avancés en matière statistique ont éprouvé de grandes difficultés à l'utiliser; et, à moins de disposer d'un personnel rompu aux problèmes de rédaction des programmes, de conduite des appareils et d'entretien de l'équipement, mieux vaut recourir à des systèmes plus simples. Les premières considérations à retenir devront être le genre de matériel disponible et la possibilité de l'utiliser au dépouillement.

57. De toute façon, on devra décider du type de matériel à utiliser et de la mesure dans laquelle on aura besoin de commis et de techniciens pour collationner les déclarations et pour préparer et revoir les tableaux. Subsidiairement, il convient d'évaluer les besoins en matériel et en personnel pour la correspondance, l'enregistrement et le suivi, ainsi que pour dresser et tenir à jour un répertoire des établissements.

Besoins en matériel et en personnel

58. Dès qu'on a décidé en général de la manière dont se fera le dépouillement, il devient nécessaire de dénombrer toutes les opérations qu'il comprend et leurs dates d'exécution. On devra calculer la besogne estimative et les besoins correspondants en matériel et en personnel, par catégorie. Ils seront coordonnés avec ceux des autres services; on commandera alors du matériel et on engagera du personnel en tant que de besoin. Pour estimer les besoins en matériel, on devra tenir compte du fait que souvent une longue attente précède la livraison. En outre, il faudra tenir compte du temps qui sera perdu à cause des pannes, ou des retards dus, par exemple, au fait que la phase précédente n'aura pas fourni suffisamment de données prêtes au dépouillement.

59. Quand le service statistique a institué un mode opératoire centralisé de conception des systèmes, on pourra tirer parti de la présence de techniciens spécialement formés pour optimiser l'exploitation. Il convient toutefois que le personnel chargé de la préparation de l'enquête revoie les méthodes proposées par les concepteurs de systèmes, car l'effet d'une opération déterminée sur la qualité des données obtenues n'est pas toujours évident.

Répartition des responsabilités

60. Etant donné que ces enquêtes sont de nature complexe et que l'exécution du plan fait intervenir différents organes du service statistique, on peut aisément négliger quelque aspect du programme. Il est

donc nécessaire d'énumérer les principales activités et de s'assurer que chaque organe est averti de ses responsabilités propres à leur sujet. Chacun devra examiner ses opérations pour déterminer dans quelle mesure sa propre tâche dépend des activités (par exemple rédaction des spécifications, préparation des plans de tableaux, choix des formats d'informatique, évaluation des séries d'essai disponibles, etc.), des autres organes et pour s'assurer que chacun sait parfaitement ce qu'on attend de lui et à quelle date.

Coordination avec d'autres enquêtes

61. Bien que le présent manuel soit conçu pour servir à organiser et à conduire des enquêtes sur la distribution et les services, on doit se souvenir que de telles enquêtes ne sont pas menées isolément. Elles s'exécutent souvent de façon conjointe avec d'autres sur les industries manufacturières, les industries minières, la construction, etc. Beaucoup d'éléments sont communs à toutes et il arrive souvent qu'une seule et même entreprise exploite plusieurs établissements dans diverses branches : il importe donc de s'assurer qu'aucune contradiction n'apparaît dans les instructions et formulaires de déclaration préparés par les différents personnels chargés d'aménager les enquêtes. De plus, pour dresser les listes d'adresses et opérer les démarches requises en cas d'abstention, une action organisée conjointement est non seulement plus efficace, mais encore aide à résoudre des problèmes de portée et à définir l'unité déclarante et l'unité statistique. Il est donc opportun de constituer dès les débuts de la planification un "comité des points communs" comprenant des représentants de chaque division et chargé d'arrêter le libellé des questions communes à tous les questionnaires. Au même comité incombera la préparation d'une liste commune d'adresses et d'un plan d'opérations subsidiaires en cas d'abstention.

62. Bien qu'il ne soit pas nécessaire de rédiger strictement dans les mêmes termes les questions communes aux différents formulaires, on ne doit pas négliger l'effet que des variantes peuvent exercer sur les autres enquêtes. Par exemple, pour définir l'unité statistique d'une manufacture, il peut y avoir lieu de se référer à des points de vente distincts qui seraient à juste raison classés comme magasins et compris dans le commerce de gros. La constitution d'un comité des points communs présente une importance particulière pour lancer les enquêtes sur la distribution et les services, lesquelles sont d'ordinaire entreprises à la suite d'enquêtes sur les industries manufacturières. On constatera, en conséquence, une tendance à négliger des questions épineuses comme la détermination des activités de vente, ce qui conduit à surestimer les activités de production.

63. Le mandat donné au comité des points communs doit comprendre des matières telles que le libellé des questions portant sur la détermination des établissements, la propriété, la forme juridique de l'entreprise, la durée de son exploitation et son emplacement. Le comité devra aussi s'assurer qu'on emploie des définitions types pour des expressions telles que "emploi", "feuille de salaire" et "dépenses d'équipement". Il sera chargé de préparer et de rédiger les instructions techniques applicables

à une tournée préalable des entreprises à établissements multiples visant à déterminer l'emplacement, la branche de classement envisagée et la taille approximative de chaque établissement exploité par l'entreprise, de façon à pouvoir envoyer le formulaire approprié. Une autre de ses fonctions consiste à déterminer le texte de l'étiquette d'expédition. Le comité peut aussi assurer la liaison avec d'autres services publics pour les questions communes à toutes les divisions, comme par exemple l'obtention de listes d'adresses ou de données.

64. De ce qui précède, il ressort clairement que les membres de ce comité doivent être des techniciens particulièrement compétents. Avec eux doivent aussi siéger des représentants des divisions d'exploitation, pour évaluer la factibilité et estimer le coût des plans dressés par le comité. Bien que le nombre des points communs soit bien moindre pour les fréquentes enquêtes courantes que pour les recensements généraux, il en importe pas moins de s'assurer dans les deux cas de la concordance des questions et définitions. Des problèmes se posent en effet car, tenus de répondre rapidement aux enquêtes courantes, les enquêtés ont quelque difficulté à préciser les données; pour les recensements généraux, ils disposent de plus de temps pour mettre au point des déclarations exactes. Il peut en résulter des disparités. Le personnel chargé de la préparation des enquêtes courantes doit donc soigneusement examiner l'action du comité des points communs.

Préparation d'un essai préalable

65. Un moyen très efficace de déterminer la factibilité et le coût d'un recensement consiste à procéder à un essai général préalable, c'est-à-dire à opérer un recensement effectif d'une partie représentative du pays - peut-être toute une province - comprenant à la fois des zones urbaines et rurales. Un tel essai doit se terminer par la publication effective des données. Il exige donc une préparation et une exploitation attentives. C'est seulement de cette façon qu'on peut être certain de prendre conscience de la plupart des problèmes susceptibles de se poser lors de l'enquête dans le pays tout entier. Inévitablement, certains problèmes, propres à une branche ou zone particulière, ne se rencontreront pas même dans un tel essai. On peut tenir pour du gaspillage le temps et l'argent consacrés à un essai préalable général; mais, par la possibilité qu'il offre de perfectionner les méthodes, de former le personnel et d'établir une bonne base d'estimation des coûts, on doit en attendre une conduite bien plus efficace du recensement et une plus grande certitude de le mener à bien.

66. Une autre forme d'essai préalable présente de l'utilité pour les problèmes relatifs à une branche déterminée : par exemple pour déterminer le libellé exact et les possibilités de réussir quand on se propose de réunir des données sur le taux d'occupation des hôtels ou sur le nombre d'entrées dans les salles de spectacles. En pareil cas, le mieux consisterait à préparer un formulaire de déclaration comprenant les rubriques objet de l'essai, lequel serait adressé à un groupe représentatif d'entreprises de la branche, avec prière de le remplir. Le personnel chargé de la préparation de l'enquête analyserait les réponses avant d'arrêter le libellé définitif des questions. Ce genre d'essai préalable peut servir

aussi bien pour les recensements généraux que pour les enquêtes courantes. Pour ces dernières, les périodes de déclaration préalable doivent être considérées comme des essais. Des plans devraient être adoptés pour un examen poussé et rapide des enquêtes initiales, de façon à pouvoir apporter promptement tout changement requis dans les méthodes ou les formulaires de déclaration.

Préparation des publications

67. La réussite d'une enquête se mesure à l'aptitude à fournir les renseignements désirés dans la limite des crédits et du temps disponibles. Avant d'entreprendre une enquête, le responsable doit connaître de façon détaillée ses objectifs en matière de données. Les usagers doivent avoir la possibilité d'examiner les résultats escomptés et le personnel d'exploitation de connaître les objectifs qu'on s'efforce d'atteindre. Le meilleur moyen d'y parvenir est que le personnel chargé de la préparation établisse un projet de tableaux qu'il soumettra à l'examen et aux observations des usagers et du personnel d'exploitation.

68. Un plan provisoire de publication, mûrement et consciencieusement préparé, forcera le personnel chargé de la préparation à se poser quelques questions épineuses mais essentielles telles que : Le questionnaire ne devient-il pas trop long et trop complexe pour satisfaire les intérêts particuliers de certains usagers ? Les questions sont-elles rédigées de sorte que les réponses ne s'excluent pas mutuellement ? Toutes les questions pourront-elles servir à des tableaux ? Un petit nombre de réponses à une rubrique particulière, en créant un risque d'indiscrétion, entraînera-t-il trop de suppressions ? Les tableaux proposés n'exigeront-ils pas de si nombreuses pages de données qu'on ne pourra pas les publier ? Peut-on combiner un un seul, par économie, les données provenant de deux ou plusieurs tableaux ? Le temps de machine nécessaire pour dresser un tableau proposé est-il excessif ? Le programme global de publication est-il tellement poussé qu'il sera trop onéreux ou trop astreignant ? Quel sera l'effet de la variance de l'échantillon sur les données publiables ? Ces questions et d'autres qui exigent des réponses se révéleront durant l'examen du plan de publication.

69. Le personnel d'exploitation doit être tenu informé de tous les tableaux proposés, de façon à pouvoir évaluer le coût et le temps de leur préparation. Cette communication doit se faire assez tôt pour qu'on découvre les grands problèmes et qu'on examine les diverses solutions. Dès qu'on a convenu d'un ensemble définitif de tableaux, le travail peut s'exécuter simultanément sur plusieurs points. Ils comprennent la définition des matières comprises dans le formulaire, l'attribution de cotes aux différentes rubriques, le codage des renseignements à perforer, la préparation des procédés de revision, la conception des procédés d'informatique pour le rassemblement et l'analyse des données et la préparation des plans concernant les tableaux. A ce stade, il est possible de préciser les estimations initiales de coût, d'effectif et de temps et d'apporter au programme toutes adaptations qui peuvent s'imposer de ce fait.

Différences de planification entre programmes de recensement et programmes d'enquête

70. La préparation des recensements généraux et des enquêtes courantes peut être confiée soit au même personnel, soit à deux personnels distincts. Ces deux modalités présentent quelques inconvénients et quelques avantages. Les principaux inconvénients des personnels distincts tiennent aux disparités qui s'établissent de ce fait entre les deux programmes et entraînent des problèmes de communication, ainsi qu'à la difficulté de chercher à conserver un personnel entraîné pendant tout l'intervalle des recensements, ce qui soulève des questions budgétaires. En revanche, avec un seul personnel pour les deux programmes, le recensement se trouve d'ordinaire négligé, car la nécessité de publier de nombreux rapports périodiques oblige le personnel à y consacrer l'essentiel de ses efforts.

71. Pour préparer les enquêtes fréquentes, on peut établir un cycle régulier d'opérations entraînant un volume de travail connu et prévisible. De cette façon, on peut déterminer l'effectif nécessaire et le coût de l'opération et être assuré de conserver le personnel qualifié disponible. Il n'en va plus ainsi s'il y a lieu de revoir l'échantillon, ce qui arrive d'ordinaire après l'exécution d'un recensement général des dossiers duquel on peut tirer des renseignements à jour sur la classification, la taille et l'appartenance. Pour y aider, on peut recourir à une partie du personnel qui a préparé le recensement et qui devient alors disponible. Il n'en résulte pas moins de fortes dépenses supplémentaires pour les opérations sur les lieux et le dépouillement.

72. L'effectif du personnel nécessaire pour dépouiller des recensements espacés varie beaucoup au cours de l'ensemble de la période, particulièrement en ce qui concerne celui chargé de l'exploitation des données et de la conduite des machines. Il peut en résulter un manque d'efficacité dans l'utilisation du matériel et des problèmes d'obtention du personnel de bureau pour la période relativement brève requise; il importe d'autant plus de dresser soigneusement les calendriers d'opérations. Quand un recensement général s'opère à la fois par correspondance et démarchage, il pose des problèmes supplémentaires liés à l'établissement de services extérieurs et à l'engagement et à la formation d'enquêteurs pour une période relativement brève. Cette opération est encore plus difficile et plus onéreuse que la création d'un service temporaire de dépouillement. Contrairement aux enquêtes qui se renouvellent et que les enquêteurs, une fois formés, peuvent assurer successivement, un recensement général demande des enquêteurs capables d'emblée, car ils disposent de très peu de temps pour profiter de la correction de leurs erreurs.

IV. GESTION ET BUDGET

Utilisation des manuels

73. Un des moyens les plus efficaces de fixer les règles qui présideront à l'exécution d'une enquête et les méthodes qui y serviront est de rédiger une série de manuels de référence qui seront distribués à tous ceux qui s'occupent de la préparation et de l'exploitation. Ces manuels doivent être conçus pour être placés dans des dossiers à fiches mobiles et porter une numérotation assez souple pour permettre de grossir chaque chapitre. Il faudra prévoir un moyen de remplacer les pages quand on modifie les instructions ou les méthodes et que le chapitre correspondant figurant dans le manuel n'est plus valable. Chaque chapitre doit porter sa date d'émission et les modifications doivent aussi être datées. Les chapitres remplacés par de nouvelles rédactions seront retirés du manuel et conservés à part pour qu'on puisse s'y reporter le cas échéant.

74. Essentiellement, le personnel chargé de la préparation rédigera les instructions prescrivant ce qu'il y a lieu de faire au personnel d'exploitation, lequel rédigera les procédés permettant d'y parvenir. Fréquemment, pourtant, pour que les buts visés soient sûrement atteints, le personnel chargé de la préparation devra rédiger les prescriptions de façon si détaillée qu'elles équivaldront à des procédés. La détermination des techniques les plus appropriées pour atteindre un objectif déterminé incombera essentiellement au personnel d'exploitation, mais les prescriptions aussi bien que les procédés définitifs doivent être arrêtés après examen des versions préliminaires par les membres des deux personnels.

75. Le manuel rédigé par le personnel chargé de préparer l'enquête doit comprendre des chapitres distincts traitant des sujets tels que :

- Etablissement de la liste d'adresses
- Expédition, réception et suivi
- Dossiers administratifs
- Opérations sur les lieux
- Sondage
- Systemes de codage
- Exploitation manuelle
- Exploitation mécanique
- Plans de publication.

Emploi d'organigrammes

76. Un autre dispositif peut accompagner le manuel ou en faire partie : la présentation des prescriptions et méthodes sous forme d'organigrammes. On peut en effet se représenter une enquête comme une série d'actions interdépendantes aboutissant à la publication de données. Chaque action découle de la précédente et conduit à la suivante. La relation qui existe entre elles peut être présentée soit sous forme de diagramme soit en langage ordinaire. Les deux modalités ont leur utilité. La présentation sous forme d'organigramme présente l'avantage de la concision et facilite à beaucoup la compréhension de la suite des tâches; elle facilite

aussi la découverte d'illogismes ou l'omission d'opérations nécessaires. On peut ainsi présenter en une seule page une documentation qui en couvrirait plusieurs en langage ordinaire.

Calendriers

77. L'une des premières démarches dans l'élaboration d'un plan de travail réaliste est la fixation d'un calendrier pour l'achèvement des activités essentielles. Etant donné que l'enquête se propose de fournir des données, la première date à fixer est celle de la sortie des publications. A partir de cette date, on remonte le temps en prenant dans l'ordre inverse toutes les opérations nécessaires. A ce stade, il suffit de déterminer les dates essentielles. Quand elles sont fixées, il y a lieu de voir les besoins à satisfaire pour respecter ce calendrier. Certaines des opérations essentielles, outre la publication, dont il y a lieu de fixer les dates, sont (dans l'ordre inverse) la préparation des tableaux, le collationnement électromécanique des données, le dépouillement des questionnaires, l'expédition des questionnaires, l'établissement de la liste d'adresses, la mise en oeuvre des méthodes de sondage, la conduite des opérations sur les lieux, l'obtention des formulaires, la détermination de leur matière et la délimitation de la portée et du champ.

78. En fixant les dates d'achèvement de chaque opération, il y a également lieu de déterminer le moment où elle doit commencer pour s'achever en temps voulu, ainsi que les autres opérations qu'il y a lieu de terminer avant de la commencer. On peut ainsi s'assurer si on a prévu un délai suffisant pour éviter de retarder les opérations ultérieures. Fixer des dates d'achèvement aux opérations essentielles, s'assurer un délai suffisant pour les terminer à temps et vérifier que les précédentes seront réalisées ponctuellement peut contribuer beaucoup au succès d'une enquête. Toutes les autres opérations que commandent ces objectifs peuvent alors s'entreprendre, ainsi que la planification des emplois attribués au personnel et au matériel. Si une quelconque de ces dates essentielles n'est pas respectée, il faudra procéder à une nouvelle évaluation des opérations ultérieures, modifier les dates ou réaffecter des ressources pour respecter le calendrier.

79. On peut dresser un tableau des opérations principales en les énumérant et en indiquant leurs dates de début et de fin dans une série de colonnes correspondant aux trimestres civils. Ce mode de présentation fait ressortir à l'évidence tout chevauchement, ainsi que les opérations dont il faut modifier le calendrier pour les achever en temps voulu. Il fournit aussi la base servant à déterminer les tâches par période et, par conséquent, à indiquer les besoins correspondants en personnel et matériel.

Calendrier d'affectation du personnel et du matériel

80. A partir du tableau susmentionné, chaque service organique chargé d'une fonction déterminée doit préparer une estimation trimestrielle du nombre et du genre de personnel et de matériel nécessaires pour atteindre l'objectif indiqué. A cet effet, il divisera les principales opérations dont il est chargé en sections pour lesquelles seront aussi indiqués les

délais prescrits. Ce procédé peut révéler les goulots d'étranglement possibles ou les opérations trop onéreuses. Quand toutes les estimations sont totalisées, on peut se représenter le tableau global des crédits et du personnel nécessaires par période. Après y avoir apporté toutes corrections jugées souhaitables, on peut dresser un tableau ferme, et tous les services connaîtront alors leurs responsabilités. Chacun devra toutefois avoir la possibilité de présenter toutes objections ou réserves qu'il croit devoir faire au sujet des objectifs qui lui sont fixés, et qui devront donner lieu à un examen approfondi avant que soient prises les décisions définitives.

Rapports d'avancement

81. Bien gérer une enquête demande qu'on sache si le travail avance conformément aux prévisions de temps et de coût. Une fois les objectifs fixés, il convient donc d'instaurer un système de rapports d'avancement.

Malheureusement, cet avancement ne se mesure pas aisément pour toutes les opérations. C'est assez facile pour des tâches comme la perforation des cartes, la revision des déclarations et les réponses à la correspondance reçue, parce que ces activités sont assez uniformes pour chaque pièce et que les tâches attribuées, la besogne accomplie et le travail restant peuvent s'exprimer par leur nombre. Toute personne qui s'y occupe peut se juger d'après son rendement et s'évaluer par comparaison avec celui de ses collègues. En outre, les taux globaux de rendement servent de base pour déterminer si le travail avance ponctuellement. En cas de besoin, on peut aisément modifier des méthodes pour accroître la productivité, ou bien licencier ou déplacer des commis peu actifs. Au contraire, des activités comme la rédaction des prescriptions applicables au dépouillement manuel pi électromécanique ne laissent pas si bien mesurer leur avancement. Il n'en reste pas moins nécessaire de fixer des délais et de suivre attentivement la marche de ces activités. Il importe particulièrement de les surveiller ainsi, car il est d'ordinaire impossible de rattraper un retard éventuel en y affectant un effectif supplémentaire, cette catégorie de personnel étant fortement spécialisée.

82. La rédaction et la mise à l'épreuve des programmes d'informatique est la tâche dont il est le plus difficile d'évaluer l'avancement. En effet, c'est seulement quand ce programme s'est révélé valable dans son application à des données réelles qu'on est certain que le travail est achevé. A toute difficulté qui se révèle au cours des derniers stades, il peut falloir longtemps pour remédier.

83. L'avancement doit se mesurer de deux façons : par le nombre d'unités produites (rapportées aux objectifs finals) et le coût correspondant. Tandis qu'on peut mesurer le nombre d'unités et le présenter en fonction du temps, les renseignements sur les coûts s'obtiennent plus difficilement et d'ordinaire trop tard pour présenter un intérêt pratique. On peut toutefois en obtenir une bonne approximation en réunissant des données sur les journées-homme ouvrées, pour chaque fonction; et en affectant à chacune de ces journées-homme un certain coût, calculé séparément pour chaque fonction. De même, on peut attribuer des coûts à des fonctions comme

l'informatique en accumulant des renseignements quotidiens sur le nombre d'heures de fonctionnement de l'ordinateur et en le multipliant par le coût horaire estimatif. Le nombre des unités produites et le montant cumulatif des coûts doivent s'analyser fréquemment par rapport aux dates d'achèvement prévues et aux crédits attribués à la tâche considérée, et les correctifs doivent être apportés assez tôt pour être efficaces.

84. Il faut prévoir au moins deux degrés dans les rapports d'avancement. Pour la direction générale, ces rapports peuvent se borner aux principales opérations et rubriques de dépenses. Pour l'encadrement subalterne, ces principales opérations doivent être subdivisées en grandes sections se rapportant à la fois aux éléments constitutifs des dépenses et aux opérations qui pourraient devenir des goulots d'étranglement, même si elles n'entraînent pas de grands frais. Il convient de ne pas oublier que le surcroît de dépenses causé par une opération qui se prolonge au-delà des prévisions affecte non seulement cette opération, mais encore réduit fréquemment à l'oisiveté le personnel et le matériel affecté aux opérations ultérieures, dont il relève ainsi le coût.

Institution d'un système de contrôle qualitatif

85. Outre respecter le calendrier des opérations, il est essentiel de maintenir une qualité acceptable. Cette considération importe particulièrement dans un recensement général, car on n'y a pas la possibilité de corriger ultérieurement une erreur, comme lors d'une enquête qui se répète chaque mois ou chaque trimestre. Pour surveiller les opérations, il convient donc d'organiser un service de contrôle qualitatif qui élabore un système de prélèvement d'échantillons dans les diverses opérations, pour permettre de vérifier si un niveau acceptable de qualité se trouve maintenu. Dans un tel échantillonnage portant sur la perforation des cartes et les opérations manuelles, le système peut être conçu pour déterminer de surcroît si les commis fournissent un rendement quantitatif acceptable. Les systèmes de contrôle qualitatif peuvent aussi être conçus pour vérifier des tâches comme la préparation de listes d'adresses par ordinateur, les opérations sur les lieux, la réponse appropriée au courrier reçu, le codage, etc. Le personnel du service chargé du contrôle qualitatif doit être indépendant : c'est-à-dire qu'il ne doit pas faire partie du service qu'il est appelé à surveiller. Avant d'être mis en oeuvre, il convient que les procédés de contrôle qualitatif soient examinés par le personnel d'exploitation et de préparation.

Maximisation de l'emploi des ressources

86. Plusieurs dispositions peuvent aider à tirer un meilleur parti des ressources. L'une d'elles consiste à utiliser judicieusement les heures supplémentaires. Souvent, un cadre responsable, désireux de respecter le délai prescrit, doit comparer les coûts respectifs de formation d'un personnel plus nombreux et d'emploi d'un moindre effectif auquel il demandera de faire des heures supplémentaires, même si le tarif en est plus élevé. Cette seconde branche de l'alternative peut se révéler moins coûteuse. Pourtant, il ne faut pas tellement recourir aux heures supplémentaires qu'elles deviennent inefficaces par suite du surmenage du personnel.

87. On peut profiter du fait que dans un recensement de nombreuses opérations se succèdent pour utiliser le même personnel aux différents stades de l'exécution. La formation et l'expérience acquises au cours d'une opération bénéficient aux suivantes et assurent un travail meilleur et plus efficace. Obtenir ces résultats exige une planification et une progression très attentives, en même temps qu'un strict respect du calendrier.

88. Un service statistique qui s'occupe à dépouiller un certain nombre d'enquêtes différentes constate de grandes variations dans l'utilisation du matériel mécanique et électronique. Un calendrier judicieux profite des périodes de ralenti et assure par là une économie générale. Il en va particulièrement ainsi en informatique.

89. L'un des principaux problèmes à résoudre au sujet des coûts tient au contrôle à exercer sur le routage des documents. La question de l'emploi optimal des commis s'y rattache directement. Il faut trouver un compromis entre une simplification de la fonction de chacun qui lui permet de l'exécuter très vite et les problèmes résultant de l'accroissement de la masse de documents transmis de l'un à l'autre. En général, doivent être confiées à un même commis toutes les opérations qu'il peut convenablement apprendre à traiter de façon efficace au cours d'une période limitée de formation. S'il faut établir des contrôles nombreux et complexes parce que trop de documents se transmettent d'un service d'exécution à un autre, l'efficacité baisse et, de surcroît, le contrôle devient difficile ou même échappe.

90. On peut contenir les coûts dans une certaine mesure, surtout lors d'un recensement général, par un judicieux recours au sondage pour les rubriques qui ne seront publiées que pour l'ensemble de la nation ou chacune de ses grandes régions. L'enquête en deviendra un peu plus complexe, car il faudra établir le plan des sondages et résoudre les problèmes de leur exploitation, mais on pourra au total réaliser une économie dans les coûts d'exécution.

91. Si on doit procéder à de massives opérations de bureau, il importe de fixer un ensemble d'étalons de mesure du rendement individuel. Ce rendement sera examiné périodiquement et signalé peut-être chaque semaine, de façon que chaque membre du personnel connaisse sa situation. Cette notation du rendement doit s'accompagner d'une notation de la qualité, car un rendement obtenu en sacrifiant la qualité se disqualifie. Un commis dont le rendement reste régulièrement inférieur aux normes doit être licencié, ou bien affecté à une autre tâche si on estime qu'il peut rendre service ailleurs. On obtient ainsi le double résultat d'éliminer les travailleurs improductifs et d'inciter les autres à rendre davantage.

92. Si bien préparée que soit l'enquête, les très nombreuses tâches qu'elle comprend et la répartition éventuelle de ses diverses phases entre les différentes divisions du service statistique suscitent fréquemment l'apparition de goulots d'étranglement, et obligent à résoudre des conflits d'opinion au sujet d'un mode d'action, de la nécessité de réaffecter du personnel, du changement des priorités, de l'adaptation des budgets, etc. C'est pour régler des situations de ce genre qu'un cadre disposant d'une

autorité suffisante doit être nommé pour coordonner la réalisation de l'enquête, comme pour prévoir et résoudre les problèmes et les différends. Bien qu'une bonne partie des renseignements dont il aura besoin pour s'acquitter efficacement de sa tâche se trouvent dans les divers rapports d'activité et rapports budgétaires, il devra aussi avoir la faculté d'inspecter personnellement toute opération et de le scruter pour trouver des solutions aux situations épineuses avant qu'elles empirent trop. Il doit avoir la faculté de réunir le personnel représentant les différentes divisions, d'adapter l'ordre des priorités et de mener avec les autres divisions des négociations visant à maintenir le programme dans les délais prescrits et dans les limites de son budget.

Différences entre recensements et enquêtes courantes

93. La répétition régulière, à bref intervalle, des enquêtes courantes, tandis qu'on doit probablement envisager tous les cinq ou dix ans seulement la réalisation d'un recensement complet de la distribution et des services, exerce un effet marqué sur le budget et la gestion. Les coûts des enquêtes qui recommencent périodiquement tendent à rester stables de mois en mois et on peut perfectionner les opérations avec le temps. Il est possible de détailler les calendriers d'exécution par jour de la semaine ou même par heure, si besoin est. Il n'en va pas de même des recensements, où la masse des travaux et des dépenses tombe au cours de l'année suivant la période couverte et où les opérations et les budgets varient beaucoup avec chacune des autres années. Dans les enquêtes courantes, l'obligation de fournir promptement et régulièrement des indicateurs essentiels oblige à respecter des calendriers très serrés que peut imposer le personnel d'exploitation, même au prix de périodes régulières et prévues d'avance d'heures supplémentaires. Dans un recensement, tout doit tendre à bien exécuter la tâche d'emblée, sinon il faut la recommencer jusqu'à correction des erreurs avant d'entreprendre l'opération suivante. Si l'erreur n'est pas décelée avant l'achèvement de plusieurs opérations additionnelles, toutes les opérations intermédiaires sont à recommencer : ce qui peut être très coûteux.

94. L'évolution du coût de la main-d'oeuvre et des autres ressources nécessaires sous l'effet de la hausse des salaires et des prix au cours de l'intervalle normal de cinq à dix ans qui sépare chaque recensement peut rapidement fausser les estimations budgétaires. La fluctuation des effectifs disponibles au cours de la même période peut aussi entraîner des difficultés. Le départ de personnel qualifié entre deux recensements est un autre inconvénient considérable, qui entraîne des modifications dans les besoins d'effectifs, non seulement de mois en mois, mais aussi de recensement à recensement. Cet inconvénient est particulièrement fâcheux quand une période de dix ans sépare les recensements.

V. PORTEE, CHAMP ET CLASSIFICATION

Introduction

95. Doivent être considérées comme parties intégrantes du présent manuel deux publications des Nations Unies : la Classification internationale type, par industrie, de toutes les branches d'activités économiques (numéro de vente E.68.XVII.8) et les Recommandations internationales pour les statistiques de la distribution et des services (numéro de vente E.75.XVIII.9).

96. La Classification internationale type par industrie (CITI) contient un exposé complet des principes suivis pour établir le système de classification et de la façon de les appliquer. L'exposé comprend des chapitres sur l'objet et la nature du système, les critères employés à son élaboration, la détermination des unités statistiques, la classification de cette unité, la structure du système et son adaptation aux conditions nationales. Les recommandations internationales concernent la portée et le champ des enquêtes; les unités à utiliser; les caractéristiques de l'unité statistique; les données (avec leurs définitions) à réunir; et un ensemble de tableaux qui présentent des recommandations sur les catégories de données à inclure dans les enquêtes espacées, les enquêtes annuelles et celles qui sont encore plus fréquentes, avec des recommandations séparées pour les pays selon qu'ils disposent déjà d'un système développé de statistiques ou s'emploient à le mettre au point dans le domaine de la distribution et des services.

97. La présentation qui suit n'est pas destinée à remplacer les publications susmentionnées, mais à faire ressortir les principaux secteurs difficiles et à proposer des méthodes pour les traiter. Il est constant que les économies des différents pays diffèrent beaucoup et que des considérations nationales conditionneront la nature des mesures à prendre. On espère toutefois qu'on cherchera à se conformer autant que possible aux recommandations, de façon à atteindre les objectifs internationaux. Chaque fois que possible, l'exposé suivant s'efforcera de faire ressortir des différences qui peuvent présenter quelque intérêt quand on considère les enquêtes annuelles et celles qui sont encore plus fréquentes par opposition aux recensements généraux.

Portée

Commerce de gros (CITI 61)

98. Les recommandations internationales proposent d'inclure toutes les catégories de grossistes dans l'enquête sur la distribution et les services. Quand on entreprend un dénombrement complet en vue de déterminer chaque établissement, cet objectif peut être atteint, bien qu'on risque de se heurter à quelques difficultés pour localiser les activités menées directement à domicile, à moins que leurs adresses figurent dans les rôles des patentes ou d'autres registres administratifs. Quand, toutefois, le plan d'enquête prévoit l'emploi combiné d'une liste d'adresses pour les grandes entreprises et d'un sondage aréolaire pour les petites (procédé courant dans

les enquêtes qui ont lieu plusieurs fois par an), apparaissent des difficultés qui exigent une modification de la portée en ce qui concerne la fraction du commerce de gros qui représente les grossistes non commerçants.

99. Plusieurs problèmes se posent. Un, le nombre de ces grossistes est relativement faible et ils ne se répartissent pas territorialement d'une manière qui se prête à un sondage aréolaire efficace. Deux, ces mêmes grossistes, en tant que représentants et courtiers, ont tendance à n'avoir que des installations réduites et travaillent fréquemment à leur domicile. Il est donc difficile non seulement de les localiser dans un dénombrement sur place, mais encore de dresser une liste d'adresses fiable. Les magasins et bureaux de vente des fabricants posent un autre problème. A moins d'opérer simultanément une enquête sur l'industrie manufacturière, il est très difficile de les reconnaître et de les dénombrer convenablement.

100. Donc, pour les enquêtes ynnuelles et celles qui ont lieu plus d'une fois par an, mieux vaut probablement se borner aux grossistes commerçants, définis par la CITI comme "les grossistes qui deviennent propriétaires des marchandises qu'ils vendent, tels que les négociants en gros ou les revendeurs, concessionnaires de produits industriels, exportateurs, importateurs, les exploitants de silos dans les gares terminus, les coopératives d'achat". Si, soutefois, on peut tenir des listes à jour de toutes les entreprises avec leur classement par genre et quelque indication de leur taille, ces listes peuvent servir soit à une enquête intégrale par correspondance, soit au choix d'un échantillon représentatif. Il est rare, toutefois, que cet ensemble de conditions soit rempli.

Commerce de détail (CITI 62)

101. Ce stade de la distribution tend à se modeler étroitement sur la répartition de la population et se prête ainsi à l'établissement d'un plan efficace de sondages aréolaires. Le problème de répartition qui se pose avec le commerce de gros n'existe pas ici et il est donc possible d'inclure tous les genres de commerce de détail dans les enquêtes annuelles ou plus fréquentes encore qui comprennent un sondage aréolaire. Etant donné le grand pourcentage des commerces de détail qui s'ouvrent et ferment en peu de temps, il est très difficile de mettre au point une liste d'adresses qui puisse servir en soi à de telles enquêtes. Une telle liste tend à ne représenter que les grandes entreprises, plus stables. Mieux vaut donc mettre au point un sondage aréolaire qui pourra servir à représenter toutes les petites entreprises et compléter ainsi la liste des grands établissements tirée d'ordinaire d'un recensement général.

Les services

102. Les services ne se prêtent pas aisément à définition. De plus, les recommandations internationales relatives à ce groupe sont fort limitatives. Elles bornent la portée des enquêtes "aux services rendus aux ménages et aux entreprises, et qui sont normalement fournis par des établissements pouvant être rangés parmi les magasins, hôtels, restaurants, théâtres, boutiques et bureaux". Il y a toutefois nombre d'exceptions, et il convient

d'étudier attentivement la portée et les caractéristiques correspondantes proposées. L'examen de la nature de ces activités montrera que certaines poseront de grands problèmes si on les inclut dans les enquêtes annuelles et les enquêtes plus fréquentes encore, surtout si on ne peut élaborer et tenir des listes d'adresses bien à jour. Dans plusieurs cas (hôtels saisonniers, production de films, émissions de radio et de télévision), les sondages aréolaires seraient inopportuns.

Champ

Généralités

103. On peut admettre qu'en pratique, il est quasi impossible d'inclure tous les commerces de gros et de détail et tous les services. Il en est ainsi parce que beaucoup de ces établissements sont exploités d'une façon qui rend leur localisation très difficile, même au cours d'un démarchage à domicile et que beaucoup ne figurent jamais sur aucun rôle fiscal ou administratif. Il serait souhaitable d'opérer quant au champ couvert certaines distinctions entre les enquêtes peu fréquentes, les annuelles et celles qui ont lieu plusieurs fois par an. Alors que les premières devraient en principe englober tous les commerces de gros et de détail et tous les services - qu'elle qu'en soit la taille, qu'ils se manifestent ou non par des signes extérieurs, qu'ils aient ou non un siège fixe -, il n'est pas pratique de vouloir le faire dans les autres enquêtes. Celles-ci, pour être réalisables, doivent se borner aux établissements à siège fixe et signes extérieurs, au moins pour le commerce de détail et les services. Imposer la condition des signes extérieurs aux commerces de gros est probablement trop restrictif, et il doit être possible d'élaborer des répertoires convenables sans procéder chaque fois à un démarchage. Les petits commerces et les services exigeront toutefois un tel démarchage des zones sondées.

104. Les principaux problèmes qui se posent tiennent : a) aux activités sans signes extérieurs, particulièrement celles qui ont lieu à domicile; b) aux établissements itinérants, ceux qui se déplacent et n'ont pas de siège fixe; et c) aux boutiques temporaires qui peuvent n'ouvrir que certains jours de la semaine ou certaines saisons de l'année. Le responsable de l'enquête devra décider si, dans le cas particulier dont il s'occupe, l'effort nécessaire pour reconnaître et dénombrer ces établissements se justifie ou s'il est fondé à recourir à des moyens tels que limites discrétionnaires de taille ou sondages aréolaires pour atténuer le problème.

Limites de taille

105. Dans la plupart des cas, les trois catégories d'établissements qu'on vient de mentionner seront certainement de petite taille. On peut donc admettre qu'une grande majorité se trouvera au-dessous d'une taille déterminée. Si on décide de les omettre de l'enquête et qu'on ne cherche pas à trouver ou à dénombrer les établissements sans signes extérieurs, itinérants ou temporaires, on peut définir l'univers statistique considéré comme formé de tous ceux supérieurs à la taille limite. Certes, quand ces petits établissements s'acquittent d'une fraction notable des activités

commerciales, l'effet sur les statistiques peut être profond. Un pays peut donc décider d'inclure dans l'enquête tous les commerces manifestes, quelle que soit leur taille. Le rapport doit l'indiquer clairement, de même que toute éventuelle limite de taille. Dans les deux cas, pour la gouverne de l'utilisateur, on s'efforcera d'évaluer l'ampleur de l'exclusion.

Sondage aréolaire

106. Dans les genres d'activités qui tendent à se répartir proportionnellement à la population, comme le commerce de détail, la restauration, les cafés et de nombreux services, on peut mettre au point un sondage aréolaire efficace qui réduira le nombre des petits établissements à contacter. Il devient ainsi possible de consacrer plus d'efforts à reconnaître et dénombrer les établissements plus difficiles à trouver, mais le recours au sondage aréolaire réduira beaucoup le volume de données sur le genre d'activité et le secteur qu'on peut publier avec un degré acceptable de variance.

Problèmes particuliers au dénombrement sur place

107. Les difficultés mentionnées jusqu'ici comme propres aux activités sans signes extérieurs, aux vendeurs itinérants et aux activités temporaires prennent une importance particulière lors d'un tel dénombrement. L'inclusion peut encore se heurter aux difficultés suivantes :

- a) Locaux fermés lors du dénombrement. Quand un commerce est fermé lors du dénombrement, l'enquêteur doit s'assurer s'il s'agit d'une fermeture temporaire ou définitive, ou bien si l'établissement n'est ouvert qu'à certaines heures ou certaines saisons. Il peut être très difficile d'obtenir des déclarations au sujet de ce genre d'établissement. Dans certains cas, comme le dénombrement dans les villégiatures ou celui des hôtels saisonniers, ce problème peut être particulièrement important.
- b) Entreprises exploitées dans des établissements d'autres branches. On peut avoir tendance à négliger les activités de distribution et les services assurés dans des établissements qui n'appartiennent pas à cette branche, tels que musées, gares, collèges, immeubles à usage d'habitation ou de bureaux, etc. Ces activités duement constatées ne devraient être énumérées et dénombrées que s'il s'agit d'exploitations privées et ouvertes au public.
- c) Comptoirs locatifs et concessions. Le risque est grand de négliger les exploitations en comptoirs locatifs ou concessions, telles que réparations horlogères dans une bijouterie, salon de beauté dans un grand magasin, banc de boucherie dans une épicerie. Cette omission peut avoir un effet marqué sur certaines catégories d'activités et peut même aboutir à une sous-estimation générale de l'ensemble des ventes au détail.
- d) Bureaux partagés. Pour réduire les dépenses d'exploitation, une entreprise loue parfois des bureaux à une autre. Ils sont ou non reconnaissables, mais le risque d'omission est grand s'il s'agit

de représentants ou courtiers. De tels arrangements peuvent représenter une fraction importante des affaires faites par les agents exportateurs ou importateurs, courtiers en grains, commissionnaires aux halles, etc.

- e) Maisons de vente par correspondance et exploitants d'appareils distributeurs. L'enquêteur peut avoir tendance à oublier ces entreprises, car leurs clients ne viennent pas sur place. Les maisons de vente par correspondance prennent contact avec eux par leurs propres catalogues ou par une publicité dans les revues et journaux; les exploitants d'appareils distributeurs traitent directement avec les établissements où ils placent leurs appareils. L'inclusion de ce type d'entreprise demande donc une attention particulière.

Problèmes particuliers au dénombrement par correspondance

108. Dans un dénombrement par correspondance, l'emploi d'un répertoire, notamment s'il est tiré de dossiers administratifs, évite certains des problèmes que pose le dénombrement sur place, mais présente quelques inconvénients propres :

- a) Distribution et services assurés par les entreprises extérieures à cette branche. Certaines des grandes entreprises comptent maints départements, dont certains dans la branche de la distribution et des services et d'autres en dehors. Par exemple, dans les dossiers administratifs, l'entreprise sera classée comme transports ferroviaires et donc considérée comme en dehors. Mais elle peut aussi posséder et exploiter des hôtels et des restaurants. Avec l'extension de l'intégration et des sociétés multinationales, il importe de reconnaître chaque organe d'une société qui rentre dans le champ de l'enquête. Avant donc le dénombrement effectif, on fera une tournée préalable de toutes les grandes entreprises pour vérifier si l'un quelconque de leurs établissements rentre dans le champ de l'enquête.
- b) Autres activités des entreprises de la branche. De façon analogue, il importe d'exclure de l'enquête les départements qui, dans les entreprises de la branche, se consacrent à des activités extérieures à celle-ci, comme par exemple une manufacture que possède un magasin à succursales multiples. Là encore, on procédera par tournée préalable des grandes entreprises.
- c) Insuffisance des dossiers administratifs. Etant donné que la compétence administrative à l'égard d'une branche peut se borner aux entreprises d'une certaine taille ou présentant des caractères particuliers, son champ ne sera pas aussi complet qu'il conviendrait aux fins d'enquête. En outre, la classification par genre d'activité peut ne pas correspondre exactement aux critères de la CITI. On étudiera donc l'effet qui en résulte sur les statistiques et, s'il est décidé d'utiliser les dossiers administratifs dans une enquête par correspondance, la publication finale devra expliquer toute limitation notable du champ.

Activités non lucratives et activités du secteur public

109. Dans maints pays, un grand nombre de transactions sont effectuées par des organismes non lucratifs ou publics. Ces organismes opèrent sur le marché d'une façon analogue aux entreprises privées : ainsi, par exemple, des coopératives agricoles écoulent du lait ou des céréales et des magasins d'Etat vendent au détail des spiritueux. Pour brosser un tableau complet de la distribution, il s'impose d'inclure ces activités à côté de celles du secteur privé.

Classification

110. Comme on l'a dit, la CITI donne un exposé très complet et très détaillé des principes de classification. En outre, elle expose les divers types d'unités statistiques et d'unités déclarantes pour lesquelles il convient de réunir des données. Reproduire cet exposé serait inutile. On mentionnera toutefois que les principes de classification et de détermination des unités statistiques et déclarantes doivent être adaptés à la situation existant dans le pays où a lieu l'enquête et que les questions qui fourniront les renseignements requis doivent être libellées avec un soin extrême.

VI. QUESTIONNAIRES ET AUTRES FORMULAIRES D'ENQUETE

111. Comme on l'a dit, les recommandations internationales sur les statistiques de la distribution et des services comprennent une présentation détaillée des données et des définitions des rubriques à inclure. Le choix de ces rubriques et des priorités se fonde sur la considération de l'utilité relative des statistiques ou de leur besoin et des difficultés de leur rassemblement. Il est constant que les besoins et les conditions varieront d'un pays à l'autre et que certaines modifications s'imposeront dans des cas particuliers. L'exposé ci-après se borne à certains des problèmes qu'on peut rencontrer pour appliquer les définitions recommandées et pour élaborer les questionnaires et autres formulaires qui contribueront à l'exécution des enquêtes.

Renseignements nécessaires au contrôle, à la correspondance et aux interpolations

112. Tout formulaire d'enquête doit demander le nom et l'adresse postale de l'entreprise propriétaire en plus du lieu de l'établissement, de façon qu'on puisse écrire à cette entreprise pour toute question qui oblige à correspondre après réception et collationnement du questionnaire dûment rempli. (Cette adresse postale diffère souvent du lieu de la situation). Si on se sert d'un répertoire pour préparer le registre d'adresses, il est très utile d'y inscrire des cotes indiquant le genre d'activité et la taille du destinataire. Ce renseignement sert à deux fins : un, il permet de couvrir retardataires et négligents; deux, il indique la mesure de l'effort de relance requis en cas d'absence ou d'insuffisance de déclaration. (Une petite entreprise n'occupant aucun salarié en mérite évidemment moins qu'une autre qui en compte 100).

113. En outre, la liste d'adresses doit comprendre pour chaque établissement un numéro matricule qui servira de référence pour l'enregistrement des déclarations, la relance des négligents, le classement des formulaires et le contrôle général.

Questions en vue de la publication et du collationnement

114. Il est évident que le formulaire doit comprendre des questions qui fournissent les données à publier. Il est moins évident qu'on doit inclure un certain nombre d'autres questions destinées à assurer que les données sont convenablement fournies et traitées. En voici quelques exemples :

- a) Si un détaillant perçoit les taxes sur les ventes et droits d'accise et que, pour l'enquête, ils doivent figurer dans le total de ses ventes, il ne suffit pas toujours, pour obtenir la réponse exacte, de demander à l'enquêté de les inclure. La bonne méthode consiste à lui prescrire de le faire, puis à lui demander s'ils ont bien été inclus et, en cas de réponse négative, de demander le montant versé au percepteur. Le reviseur dispose alors de tous les renseignements nécessaires pour parvenir à la réponse exacte.

- b) Fréquemment, il arrive qu'en un même lieu et sous une même raison sociale opère plus d'une entreprise; c'est, par exemple, le cas d'un grand magasin ayant concédé l'exploitation d'un rayon parfumerie ou chaussures. Il est nécessaire alors de rédiger une question qui fournisse ce genre de renseignement et d'obtenir des indications suffisantes pour permettre de traiter convenablement la question.
- c) Inclure des questions "de recoupement" peut se révéler très utile. Par exemple, on demandera à l'enquêté de comparer ses recettes et les salaires de son personnel pour voir s'ils se concilient. On peut encore lui demander de vérifier la concordance des totaux de plus d'une rubrique : peut-être du chiffre global des ventes et de leur répartition par marchandise. De tels contrôles doivent rester simples de façon à pouvoir s'exécuter rapidement.
- d) Pour les enquêtes annuelles ou moins fréquentes, les recommandations internationales préconisent des déclarations pour l'année civile. Toutefois, si la plupart des établissements utilisent habituellement l'année fiscale, on peut réunir uniformément des données sur cette base. Tout en indiquant dans le questionnaire la période préférée, il sera nécessaire de demander à l'enquêté celle qu'il a choisie et, s'il s'agit de l'exercice fiscal, d'en préciser les dates, car toutes les entreprises n'utilisent pas le même. Les renseignements portant sur une période autre que celle prescrite ne devront être acceptés que compte tenu de leur influence sur les résultats : si cette influence est notable, on s'efforcera d'obtenir pour la période requise, des renseignements essentiels, comme l'emploi et les ventes qui pourront servir à calculer au prorata les autres rubriques importantes.
- e) Un problème particulier se pose dans les enquêtes qui ont lieu plusieurs fois par an, quand l'entreprise ne fonctionne qu'une partie de l'année, ce qui est le cas des hôtels saisonniers ou des établissements sis dans les villégiatures. Il est souvent impossible de distinguer ces cas de celui d'une entreprise qui a cessé toute activité ou quitté ses locaux, si l'enquêteur n'obtient des réponses à toute une série de questions. L'omission des établissements saisonniers ou temporairement fermés risque d'entraîner une forte distorsion réductrice.

Distinction entre réponses omises et réponses "néant"

115. Un problème se pose couramment dans le dépouillement des questionnaires quand un enquêté saute une question dont la réponse pourrait être "nul" ou "néant". L'enquêteur se trouve alors devant un dilemme : déterminer s'il y a abstention ou si l'enquêté a jugé la question sans objet. Bien qu'il soit utile de prescrire de répondre à toutes les questions, le cas échéant par les mentions "néant" ou "nul" ou encore par un tiret, beaucoup d'enquêtés ne s'y conformeront pas. On peut réduire la portée de ce

problème en posant une question appelant une simple réponse par oui ou par non, qui servira à indiquer s'il faut porter la mention "néant" ou la traiter comme ayant été omise.

Rédaction de la question sur le genre d'activité

116. La présentation des données par genre d'activité est l'un des objectifs essentiels des enquêtes sur la distribution et les services. Parvenir à une classification appropriée conforme aux normes établies par la CITI exige l'application de certains critères objectifs qui, le plus souvent, demandent qu'on obtienne des réponses à un ensemble type de questions. Pour certains genres d'activités, comme celles des coiffeurs, des émetteurs de radio ou des producteurs de films, l'appellation peut suffire, mais dans la plupart des cas il est nécessaire d'approfondir davantage pour classer convenablement. Par exemple, la mention "habillement" ne suffit pas, car il peut s'agir de la fabrication, de la vente en gros ou au détail ou de la réparation. De plus, l'enquêté peut se considérer comme un détaillant, alors que, pour la CITI, c'est un grossiste. La rubrique genre d'activité doit donc comprendre des questions qui permettent de classer l'établissement d'abord comme commercial, manufacturier ou des services. S'il est commercial, s'occupe-t-il essentiellement de gros ou de détail ? S'il fournit des services, par exemple des réparations, s'agit-il de son activité principale ? S'il s'occupe de commerce et si le programme de publication prévoit des tableaux correspondant à une répartition détaillée des activités, il faut poser des questions supplémentaires sur la répartition des ventes par genre de marchandise.

117. Si on dispose d'une source permettant une classification préliminaire, on peut poser sur le genre d'activité une question libellée en fonction du type général d'activité de l'établissement considéré. Ainsi, on inclura une question différente pour les entreprises de l'alimentation, celles de l'habillement, etc. On peut ainsi poser des questions pertinentes, avec la terminologie appropriée, rendant le questionnaire plus intelligible. On peut aussi classer selon une répartition plus détaillée des activités et poser des questions précises appropriées au genre général d'activité, comme par exemple sur la vente de denrées congelées dans les magasins d'alimentation.

Détection des déclarations communes

118. Vu la difficulté de connaître toutes les entreprises à établissements multiples avant de procéder à une enquête par correspondance, on a besoin d'un moyen de les découvrir lors du dénombrement. A cet effet, le formulaire doit demander la liste de tous les établissements de l'entreprise et les renseignements de base correspondant à chacun, tels que genre d'activité (autodésignation), volume des ventes, salaires annuels et effectif du personnel. Comme on peut présumer que les grandes entreprises auront été déterminées par avance et invitées à fournir des données séparées pour chaque établissement, le nombre et la taille de celles ajoutées lors du dénombrement seront faibles.

Codage du formulaire

119. Une cote doit être attribuée à chaque rubrique ou autre genre d'information qui sera mis sous forme de tableau, ou utilisé au tri, ou requis pour identifier l'établissement. Cette cote peut varier selon le genre d'appareils utilisés, tels que machines comptables, matériel à cartes perforées ou ordinateurs. Le choix judicieux des cotes peut influencer de façon positive sur le coût et la qualité du travail et doit se faire en connaissance des modalités d'exploitation des différentes rubriques. Par exemple, avec des machines à cartes perforées, on veillera dans toute la mesure du possible à porter sur la même carte les données qui doivent figurer dans le même tableau.

Importance du précodage

120. On économisera beaucoup de temps et d'argent en examinant chaque rubrique du formulaire pour voir s'il est possible de libeller les questions de façon à pouvoir imprimer directement sur le formulaire les cotes préalablement attribuées, au lieu de les faire porter par un commis. On s'assure ainsi un double avantage : le commis n'aura pas besoin de chercher et d'inscrire la cote; et il sera plus facile au perforateur de la trouver et de la lire si elle est imprimée que si elle est manuscrite et peut-être illisible.

Rédaction des instructions

121. Quand le questionnaire est rempli par l'enquêteur, un manuel contient ses instructions et on peut consacrer le temps nécessaire à le former et à s'assurer qu'il les comprend. Si, au contraire, l'enquêteur doit laisser le questionnaire à l'enquêté ou si l'enquête se fait par la poste, c'est par le moyen des instructions portées sur le formulaire ou jointes qu'on devra obtenir les renseignements requis. Quand on combine le dénombrement sur place et la correspondance, on risque une distorsion si les instructions suivies par l'enquêteur diffèrent de celles figurant sur le formulaire posté. Quand des services différents s'occupent du recensement et des enquêtes annuelles ou plus fréquentes encore, on devra veiller spécialement à bien coordonner les instructions. Sinon, on risque fort que les deux séries de données ne soient pas comparables.

122. Les instructions destinées aux enquêtés doivent comprendre de façon précise et complète tout ce qui est nécessaire pour expliquer ce qu'on attend d'eux et pour éviter tout malentendu. Des instructions longues et diffuses ont moins de chances d'être lues et comprises que celles qui sont brèves et pertinentes. Si de longues instructions s'imposent toutefois, il est courant d'en faire figurer les points essentiels sur le questionnaire même et le reste sur une feuille complémentaire. Des entretiens préalables à l'enquête avec des destinataires envisagés fournissent souvent de précieuses indications sur les points qu'il y a lieu de faire ressortir dans les instructions.

123. Voici une liste récapitulative des points figurant normalement dans les instructions imprimées sur les questionnaires ou sur des feuilles complémentaires :

1. Base légale de l'enquête.
2. Caractère confidentiel des déclarations.
3. Genre d'établissements invités à répondre.
4. Acceptabilité des estimations à défaut de livres comptables.
5. Date de remise du questionnaire complété.
6. Mode d'envoi du questionnaire complété.
7. Période couverte par l'enquête et règles relatives à l'inclusion des données qui se rapportent à l'exercice fiscal.
8. Définitions des diverses rubriques demandées.
9. Instructions sur l'établissement d'indications chiffrées telles que taux d'occupation des hôtels ou de fréquentation des salles de spectacle.
10. Instructions propres à des branches particulières où on s'attend à des problèmes pour de nombreux enquêtés, comme dans le cas de location de comptoirs dans les grands magasins.

Problèmes résultant de la diversité des activités

124. La grande diversité des activités comprises dans le secteur de la distribution et des services crée des difficultés, car obtenir une classification appropriée oblige à poser nombre de questions, dont certaines peuvent ne concerner qu'une faible fraction des enquêtés. En outre, des questions spéciales ne s'appliquent qu'à des genres d'activités déterminés. Si on n'utilisait qu'un seul modèle de formulaire, sa longueur et sa complexité embarrasseraient et irriteraient les enquêtés et réduiraient probablement beaucoup la proportion des réponses. Les solutions apportées à ce problème doivent différer selon qu'il s'agit d'enquêtes par entretiens sur les lieux ou par correspondance. On considère toutefois qu'on aura rédigé un certain nombre de questionnaires différents, à raison d'un par catégorie d'établissements auxquels seront posées les mêmes questions, par exemple magasins d'alimentation, magasins de vêtements, spectacles et divertissements, grossistes pétroliers, etc.

125. Dans les enquêtes par entretiens sur place (menées à l'occasion de recensements généraux ou dans les zones sondées), l'entretien initial doit comprendre une série de questions visant à déterminer s'il s'agit d'un établissement manufacturier, commerçant (gros et détail) ou des services et à choisir le questionnaire correspondant à la catégorie. L'enquêteur peut alors laisser le questionnaire pour qu'il soit rempli et renvoyé au bureau de statistique, ou bien la liste d'adresses permettra à ce bureau d'envoyer à l'établissement le questionnaire qui convient.

126. Dans les enquêtes par correspondance, il sera nécessaire de trouver dans les registres d'expédition assez d'indications pour envoyer le questionnaire approprié. Les formulaires seront conçus en fonction du genre d'indications disponibles. Par exemple, si les registres ne permettent pas de distinguer entre commerce de gros et de détail, le questionnaire devra porter sur les deux. Ainsi, le nombre des différents formulaires et leur complexité dépendent des indications dont on dispose pour les expédier. Un problème particulier se pose quand le dossier servant à cette expédition comprend quelques entreprises qui échappent à toute classification. En pareil cas, il faudra élaborer un formulaire spécial qui non seulement convienne aux établissements de la branche, mais encore qui fasse reconnaître ceux qui sont en dehors. Toutefois, les données ne seront demandées qu'aux établissements de la branche. Une nouvelle expédition pourra aussi se révéler nécessaire pour les établissements dont on n'a obtenu d'abord que des données insuffisantes; tel sera le cas si le formulaire utilisé pour le contact initial est rendu trop long et trop complexe par l'inclusion de questions particulières concernant tous les genres d'activités. Le second contact peut se borner aux établissements assez importants pour exercer une influence marquée sur les données. Pour les autres, on pourra interpoler les données.

Conception des formulaires

127. Les formulaires d'enquête sont les documents essentiels d'obtention de toutes les données qui doivent être traitées pour préparer les tableaux et publier les résultats des enquêtes. Leur disposition, leur présentation et leur contenu exerceront une influence marquée sur la qualité de ces résultats, le coût de la réunion, du collationnement et de la mise en tableaux, et la rapidité de la publication. Un libellé ambigu, des erreurs d'impression ou même un mauvais espacement du texte ont amené certains pays à abandonner d'importantes enquêtes ou à consacrer des ressources rares à s'efforcer de corriger les conséquences de ces fautes. Dans l'autodénombrement, où l'enquêté remplit le formulaire sans l'aide de l'enquêteur, il est arrivé que même un caractère d'imprimerie mal choisi provoque un malentendu généralisé. Il peut aussi en aller de même quand un titre, important pour l'intelligence des mentions qui le suivent, ne se détache pas suffisamment et que, de ce fait, l'enquêté le saute. Un essai préalable des formulaires fait souvent ressortir ces défauts à temps pour qu'on les corrige. Cet essai n'a pas besoin d'être poussé ni scientifiquement conçu. Si on propose divers formats, il permet de connaître le plus indiqué.

128. Une bonne présentation tient aussi à l'utilisation appropriée du format choisi. Trop souvent, la place prévue pour répondre à une question ne suffit pas, ce qui donne à l'enquêté le sentiment que l'organisme responsable n'y connaît rien. En revanche, un espace excessif allonge le questionnaire plus que de besoin, portant l'enquêté à s'imaginer que le remplir demandera plus de travail qu'il n'en faut en réalité. Si on s'attend que, sur de nombreux questionnaires, les réponses soient dactylographiées, l'intervalle entre les lignes doit correspondre à celui des machines à écrire. En outre, on devra tenir compte des cadres nécessaires à l'administration. Les mentions qui y seront portées rentrent dans trois catégories générales :

- a) Données de contrôle, telles que numéro de l'établissement, identification de ceux qui dépouillent les documents et dates des diverses opérations;
- b) Cotes, telles que celles indiquant l'emplacement, le genre d'activité, la taille de l'établissement et le mode de propriété. Ces cotes ne sont pas nécessaires si les questions sont conçues pour faire l'objet d'un codage préalable;
- c) Chiffres obtenus par le calcul, tels que l'effectif moyen occupé, les gains moyens, etc.

129. Vu la nécessité de changer fréquemment les noms et adresses, la case correspondante doit laisser une place suffisante à cet effet. En outre, chaque questionnaire doit comprendre un espace approprié où l'enquêté inscrira les observations nécessaires pour expliquer ses réponses le cas échéant.

Succession et libellé des questions

130. Autant que possible, il est bon de placer en tête les questions les plus faciles. Si l'enquêté se décourage dès qu'il commence à remplir le questionnaire, il sera porté davantage à l'écartier. De même, les questions doivent être disposées de façon à rapprocher celles qui découlent naturellement l'une de l'autre. Quand figurent certaines questions propres à une partie seulement des destinataires, elles doivent se trouver à la fin. Chaque rubrique doit porter un numéro qui facilitera sa mention dans les instructions et dans toute correspondance avec l'enquêté, de même que dans la rédaction des méthodes administratives de dépouillement des réponses. Si l'on se sert de plus d'un modèle de formulaire, on s'efforcera par tous les moyens de garder le même numéro aux mêmes rubriques (par exemple salaire et emploi), même si certains formulaires doivent de ce fait comprendre une rubrique qui n'appelle pas de réponse et porte la mention "sans objet". La rédaction des règles de collationnement et d'administration s'en trouvera grandement simplifiée.

131. Certes, les instructions présentent de l'importance et doivent être soigneusement rédigées, mais le libellé des questions exige encore plus de soin. En les rédigeant de façon facile à comprendre par les enquêtés et les enquêteurs, on favorisera beaucoup des réponses promptes, complètes et sûres. Certains enquêtés omettront de lire des instructions séparées, mais on peut présumer que tous liront les questions avant d'y répondre. Tout doit donc être fait pour leur donner une teneur claire et concise et pour éviter toute ambiguïté. Si une question se prête à deux interprétations, également probables, dont une seule est la bonne, la moitié des réponses sera fausse.

132. Pour certaines rubriques, mieux vaut demander à l'enquêté de répondre en cochant une case. Ce procédé pourra servir à obtenir des renseignements sur la catégorie de clientèle, la méthode de distribution des produits, la forme juridique de l'entreprise, etc. Il peut aussi servir à obtenir des données quantitatives si on peut accepter des réponses données par classes de taille. Pour prévenir toute omission, il est bon d'inclure une catégorie

"divers", obligeant l'enquêté à la cocher s'il estime qu'aucune des autres cases ne convient. Un commis doit examiner ensuite ces marques pour déterminer la case qu'il aurait fallu cocher. Ce mode de réponse offre encore l'avantage de permettre d'imprimer à côté de chaque case la cote qui lui a été préalablement attribuée, ce qui évite le codage manuscrit.

133. Enfin, on doit mentionner que l'emploi de la couleur et du grisé peut améliorer les réponses, les opérations de perforation et l'exploitation en général, en montrant clairement à la personne chargée du dépouillement les endroits où doivent figurer des réponses et en faisant clairement ressortir les blancs qui n'ont pas été remplis. Des signes tels que flèches et pointillés sont encore utiles à cet égard. Il peut aussi y avoir intérêt, pour les mentions de grandeur, à réserver une case par chiffre, car ainsi l'enquêté apporte plus de soin à répondre et le dépouillement se trouve facilité. Une inscription judicieuse peut aussi faire arrondir les chiffres, ce qui réduit les risques d'erreur courus en arrondissant ensuite et supprime une opération.

VII. REUNION DES DONNEES

Observations générales

134. Pour choisir le système de collecte des données, il y a lieu d'analyser les problèmes spéciaux que présente le champ à couvrir, le genre de renseignements à réunir et les ressources disponibles en crédits, personnel qualifié et matériel spécialisé. Plusieurs caractéristiques devraient rester présentes à l'esprit pour concevoir des enquêtes sur la distribution et les services. De nombreuses activités de cette branche comptent un grand nombre de petits établissements. Souvent, leurs propriétaires sont illettrés et éprouvent de grandes difficultés à comprendre, même en cas de dénombrement personnel. En outre, ils ont tendance à ne pas tenir de livres : les données pour la période considérée - année, trimestre ou mois - sont donc des estimations d'ordinaire d'autant moins bonnes que cette période est plus longue. Fréquemment, leurs activités soit ont lieu à domicile ou dans des locaux sans signes extérieurs, soit sont nomades. Ils ont tendance à se désintéresser des demandes par correspondance. Le grand nombre des établissements rend onéreux un dénombrement personnel.

135. Dans le plan d'enquête, il importe aussi de considérer que, surtout dans le commerce de détail, une bonne partie de l'activité provient d'établissements appartenant à des entreprises qui en comptent plusieurs. Une telle entreprise rassemble d'ordinaire à son siège central les données concernant chacun de ses établissements : mieux vaut donc demander les renseignements à ce siège que d'envoyer à chaque établissement (qui peut même ne pas disposer de toute la documentation nécessaire) des enquêteurs ou des questionnaires. Si l'on est en mesure de dresser assez tôt une liste de la plupart de ces entreprises en préparant les grands recensements, on pourra en faire une tournée préalable pour obtenir une liste de tous les établissements (y compris auxiliaires) où figureront assez d'indications de classement pour qu'on puisse choisir le bon formulaire lors de l'enquête effective. Tous ces formulaires peuvent alors être envoyés au siège central. Ce procédé a l'avantage de charger l'entreprise de coordonner les déclarations provenant de ses divers établissements et de concilier en cas de divergence les données fournies par chacun. De toute façon, le siège central préférera souvent s'assurer des réponses et transmettre les questionnaires dûment remplis. Toutefois, en procédant ainsi, on devra veiller à instaurer des contrôles pour prévenir tout double emploi soit de la part d'enquêteurs, soit par l'envoi direct des formulaires aux établissements. Les enquêteurs doivent recevoir une liste d'entreprises ou d'établissements déterminés à ne pas dénombrer; et toute double inscription doit disparaître des listes d'adresses. Quand des enquêtes ont lieu à la suite d'un recensement général, on peut tirer de ce recensement des listes d'entreprises à succursales multiples qui indiqueront aux enquêteurs celles qu'il y a lieu d'omettre.

136. Le plan d'enquête dépend d'un autre facteur : la grande variété des activités et genres d'établissements de la distribution et des services. Les commerces de gros et de détail se définissent par le transfert de marchandises, mais des renseignements sur la catégorie de clientèle sont

essentiels pour déterminer le genre d'activité. De plus, une répartition détaillée par activité se détermine selon diverses combinaisons des genres de marchandises vendues. Dans les services, la gamme des activités est extrêmement étendue et comprend la fourniture de repas, la rédaction de textes publicitaires, la production de films cinématographiques, la location de machines, la réparation de pendules, etc. Cette extension oblige à concevoir le système pour pouvoir choisir le bon formulaire si on procède par correspondance, ou à former soigneusement l'enquêteur à déterminer exactement le genre d'activité. Il faut également prévoir la détection et la correction des erreurs quand on s'est trompé de formulaire.

Méthodes de dénombrement

137. Trois méthodes fondamentales s'offrent pour obtenir les données recherchées dans les enquêtes. La première est d'envoyer des enquêteurs dans chaque établissement. Si ce dernier est petit et que le questionnaire soit assez simple pour permettre d'obtenir les réponses lors de l'entretien, l'enquêteur peut le remplir à ce moment. Si, pour une raison quelconque, ce n'est pas possible, soit que la personne compétente ne soit pas disponible ou que les données requises soient trop complexes, l'enquêteur peut laisser le questionnaire qu'on renverra par la poste ou qu'il repassera prendre.

138. La seconde méthode consiste à envoyer le questionnaire directement à l'enquêté. Elle ne peut servir que s'il est possible d'élaborer un répertoire des établissements et de s'y fier et que si le service postal marche convenablement. Les questionnaires dûment remplis peuvent être retournés par la poste ou ramassés par des enquêteurs.

139. La troisième méthode s'emploie quand un pays a mis au point un système suffisamment perfectionné de registres administratifs, grâce auxquels les données recherchées dans l'enquête sont réunies au cours de diverses fonctions de l'administration nationale, concernant par exemple l'impôt sur le revenu, la sécurité sociale, l'assurance-chômage, etc., et que le système s'assortit d'une méthode d'identification certaine des établissements.

140. En pratique, les trois méthodes peuvent se combiner, selon les conditions particulières qui règnent à l'époque de l'enquête dans le pays considéré. Ainsi, il peut être possible d'élaborer un répertoire des grandes entreprises auxquelles les formulaires seront adressés directement et de le remettre pour leur indiquer les entreprises à exclure, aux enquêteurs chargés de ramasser les questionnaires.

Dénombrement sur place

141. Pour les pays qui envisagent d'exécuter pour la première fois un recensement général de la distribution et des services, il serait particulièrement utile de procéder à un dénombrement complet servant à dresser un registre de tout établissement reconnu au cours de l'opération, qu'il appartienne ou non à la branche. Un tel répertoire peut se révéler très précieux pour fixer les bases de sondage. Un dénombrement sur place complet s'impose aussi quand il n'est pas possible d'élaborer un répertoire

auquel on puisse se fier pour l'expédition des questionnaires. Dans un tel dénombrement, on peut aussi recourir à un sondage aréolaire pour certains secteurs nationaux de faible activité économique, pourvu que la délimitation de ces aires suffise à permettre de les identifier exactement. De toute façon, le dénombrement sur place exige un gros effort pour établir un jeu détaillé de cartes géographiques qui serviront à contrôler les affectations données aux enquêteurs.

142. Bien qu'un dénombrement sur place de la distribution et des services présente de nombreux caractères communs avec un recensement de la population, certaines différences rendent ce dénombrement bien plus difficile et exigent de plus grandes qualités de la part des enquêteurs. Incidemment, on peut combiner le recensement de la population avec celui de la distribution et des services et, en théorie, réduire considérablement les dépenses de cartographie, de formation professionnelle et les frais généraux. Toutefois, la complexité qui en résulte, ainsi que les différents délais d'exécution requis, embarrassent beaucoup l'opération et peuvent se solder par un double échec.

143. Le dénombrement sur place des établissements de la distribution et des services est très complexe et oblige à instaurer des contrôles très stricts comme à dispenser une formation poussée aux contrôleurs et enquêteurs. S'il n'est pas indispensable d'utiliser un personnel ayant des connaissances économiques, son niveau doit être tel qu'il comprenne les problèmes de classement. S'il ne peut assimiler la notion de portée ou s'il n'a pas le sens des chiffres, de grosses difficultés peuvent s'ensuivre. On devra prendre le temps de former convenablement les enquêteurs tant aux techniques de dénombrement qu'aux objectifs de l'enquête. Ils doivent être conscients des conséquences de leurs actions et omissions. A maintes reprises, il s'est révélé nécessaire de dénombrer derechef d'importants secteurs d'un recensement parce que les instructions avaient été mal comprises. Si les crédits disponibles sont limités, mieux vaut réduire la portée de l'enquête (et, partant, l'effectif des enquêteurs) que diminuer la qualité du dénombrement.

144. Pour rédiger les manuels de formation et organiser les cours, il est opportun de confier la plus grande partie possible de la tâche à quelques membres du personnel de préparation. En pratique toutefois, étant donné les gros effectifs du personnel à former et l'étendue souvent considérable des territoires, il est nécessaire de former des groupes intermédiaires qui, à leur tour, formeront les enquêteurs. Ce procédé devra être aussi limité que possible, car il en résulte une dilution et une déformation de la documentation originale. Un bon moyen consiste à former d'abord ceux qui deviendront ensuite contrôleurs et de s'assurer, par des entretiens organisés par les cadres chargés de la planification de l'enquête, qu'ils ont bien assimilé la matière. Ces contrôleurs, à leur tour, procéderont de même pour s'assurer des connaissances des enquêteurs. Les exercices donnés en classe ne sauraient remplacer la mise en pratique pour indiquer à quel point la matière a été assimilée. Même si l'instruction a été bien préparée, on ne devra jamais admettre qu'il suffit de la donner aux enquêteurs pour qu'ils accomplissent un bon travail. Il faudra sans tarder examiner ce travail de façon à pouvoir corriger des erreurs éventuelles avant qu'elles causent trop de dommage. L'idéal serait qu'à cet effet, le contrôleur refasse lui-même le dénombrement d'un district.

145. On tire un meilleur parti de la formation quand on emploie les mêmes enquêteurs à des enquêtes périodiques, par exemple mensuelles ou trimestrielles. Toutefois, même en pareil cas, il est bon d'élaborer un programme régulier de vérification périodique des pratiques suivies. Il peut se révéler difficile de trouver un bon personnel d'enquêteurs pour un recensement général, car la période d'emploi est très brève (quoique plus longue que pour un recensement démographique). Etant donné que cette phase de l'opération peut absorber le gros des crédits, on peut avoir tendance à offrir des salaires inférieurs, ce qui attire d'autant moins les auxiliaires compétents. S'il existe un programme annuel, sans programme mensuel ou trimestriel, la formation des enquêteurs devient très onéreuse, d'autant qu'ils risquent de se renouveler largement d'une enquête à la suivante, puisqu'un long intervalle les sépare et que les enquêteurs ont tendance à prendre d'autres emplois.

146. Après avoir suivi les cours d'instruction, l'enquêteur doit porter avec lui lors du dénombrement un manuel qui traite de toute la matière qu'il doit connaître pour opérer avec succès. Doivent y figurer notamment toutes les questions de méthode qui l'intéressent, ainsi que les principaux problèmes auxquels s'attendre. Voici une table des matières typique :

- a) Réglementation. Obligations pour les enquêteurs de fournir des déclarations exactes et de tenir pour confidentielles les informations données par les enquêtés. Interdictions imposées aux enquêteurs, telles que prospection ou vente.
- b) Extraits de la loi ou ordonnance prescrivant le recensement.
- c) Méthodes. Organisation du personnel d'enquête. Comment obtenir des réponses et transmettre les questionnaires dûment remplis. Que faire si l'enquêté veut qu'on lui laisse le questionnaire pour le remplir ou pour l'expédier au service statistique. Que faire si l'enquêté refuse de répondre. Comment signaler les établissements qui ont changé de propriétaires au cours de l'année. Comment rédiger les rapports d'activité.
- d) Liste des établissements que l'enquêteur ne doit pas visiter. Elle peut comprendre ceux des entreprises à succursales multiples que le service statistique central contactera directement.
- e) Portée du recensement. Définition de l'établissement et genres d'établissements à inclure. Genres d'établissements courants à ne pas inclure.
- f) Liste des questionnaires et autres formulaires servant au recensement et bref exposé du contenu, de l'objet et de la portée de chacun.
- g) Instructions détaillées pour chaque question du questionnaire.
- h) Instructions particulières pour les genres déterminés d'établissements ou de branches d'activité. Comment traiter les activités de caractère mixte qui se rencontrent fréquemment

(par exemple manufacture et commerce de gros, commerce de détail et services, comptoirs en location).

- i) Entreprises à succursales multiples ne figurant pas sur la liste précitée en d). Comment reconnaître les établissements ou sièges centraux des entreprises à succursales multiples.
- j) Index alphabétique des matières contenues dans le manuel.

147. Pour préparer le dénombrement, chaque partie du pays aura été divisée en districts. Pour les enquêtes par sondages aréolaires, les zones choisies auront été divisées de même. Ultérieurement, on dressera pour chaque district une carte dont l'enquêteur se servira pour dénombrer ceux auxquels il est affecté.

148. L'enquêteur est tenu d'inscrire les renseignements relatifs à chaque établissement en les portant sur un registre contenant toute la documentation nécessaire au contrôle de l'exécution du dénombrement. Pour chaque établissement, il devra déterminer s'il rentre ou non dans le champ de l'enquête. S'il y rentre, il faut l'inclure et s'entretenir avec le responsable. Pour réduire le coût, il peut être opportun de prescrire à l'enquêteur de ne pas inscrire des locaux professionnels manifestement étrangers à l'enquête, tels que cabinets de médecin ou d'avocat, bâtiments agricoles, etc. Toutefois, on soulignera qu'il importe d'enregistrer et d'inscrire tout établissement donnant lieu à un doute.

149. Pour contrôler le dénombrement, pour déterminer si l'établissement rentre dans l'enquête, pour choisir le formulaire correspondant et pour pouvoir reconnaître exactement l'établissement, il est nécessaire de savoir :

1. Le nom et l'adresse de l'établissement.
2. Les éléments de classification qui déterminent s'il rentre dans le champ de l'enquête.
3. Si l'établissement appartient à une entreprise à succursales multiples dont les déclarations seront fournies par le siège.
4. Si un questionnaire a été obtenu, s'il sera envoyé directement au service statistique central ou si l'enquêteur doit revenir le ramasser.
5. La date de distribution du questionnaire.
6. Le nom de la personne interrogée.
7. Le numéro matricule de l'établissement, aux fins de contrôle.

150. Etant donné que les districts de dénombrement sont délimités d'avance, il est possible d'en tracer le contour de façon à les contenir dans les divisions géographiques pour lesquelles on désire connaître ou publier

des données. Donc, si on connaît le numéro du district, tous les établissements qui portent ce numéro peuvent faire l'objet d'un codage géographique par provinces, villes ou autres secteurs.

Constitution et utilisation d'un répertoire

151. L'idéal serait qu'avant de mener une enquête, on dispose d'un répertoire de toutes les unités statistiques à y inclure, avec des renseignements sur le lieu, le genre d'activité et la taille. Par là se trouve défini l'univers statistique à dénombrer; les informations sur le genre d'activité peuvent servir à préparer les questionnaires appropriés, celles sur le lieu à opérer la répartition géographique et celles sur la taille à prendre les décisions requises en matière de sondage et de champ couvert. Le répertoire permet de contrôler le dénombrement, qu'il ait lieu par poste ou sur place, et fournit un moyen d'interpoler les données à défaut de réponses. Comme un tel répertoire ne peut être complet qu'après achèvement de l'enquête, celui qui la dirige doit se proposer d'en approcher le plus possible avant.

152. Au minimum, le répertoire doit comprendre les renseignements suivants :

1. Nom et lieu de chaque établissement.
2. Adresse pour la correspondance, qui peut différer du lieu de l'établissement.
3. Nom et adresse du bureau ou siège central pour les établissements appartenant à des entreprises à succursales multiples.
4. Genre d'activité, description ou cote.
5. Cote de taille, jugée normalement d'après l'effectif occupé, mais éventuellement sur le volume des ventes ou le montant des recettes.
6. Source et date des renseignements.

153. Etant donné que la distribution et les services se caractérisent par un grand nombre de petits établissements, qui en outre changent souvent de main, la création et la tenue à jour d'un répertoire complet peuvent se révéler difficiles et onéreuses. En conséquence, certains pays fixent une limite de taille et n'inscrivent dans le répertoire que les établissements qui sont en dessous. Le procédé est acceptable quand d'autres dispositions seront prises pour couvrir les autres établissements : par exemple en dresser sur place une liste complète ou se servir d'autres sources de données comme les registres administratifs ou fiscaux. En revanche, si l'enquête doit se fonder sur un répertoire qui ne retient que les établissements d'une taille supérieure à la limite fixée, on ne sera jamais certain de les avoir tous inclus. Certes, si un pays a élaboré un moyen de tenir à jour un registre de toutes les entreprises à des fins administratives ou fiscales, on peut en tirer les renseignements nécessaires au répertoire considéré.

154. L'utilité d'un répertoire dépend du genre de données qu'il contient, de son exhaustivité et de son exactitude. Pour l'établir, on peut puiser à différentes sources, dont la plupart présentent des défauts. Chacune doit être examinée soigneusement avant usage, et on doit veiller à remédier à ses défauts. En même temps, il peut être nécessaire de sacrifier quelque peu l'exhaustivité ou l'exactitude pour que le coût reste acceptable, mais de telles décisions doivent être prises sciemment, et on doit chercher à mesurer et à décrire les défauts. On trouvera ci-après quelques-unes de ces sources pour les répertoires de la distribution et des services :

- a) Liste complète établie sur les lieux. C'est probablement la façon la plus onéreuse de réunir les renseignements nécessaires à un répertoire. Toutefois, surtout dans un pays qui lance un programme de statistiques économiques, c'est incontestablement la plus efficace. Des enquêteurs entraînés peuvent rechercher toutes les exploitations manifestement reconnaissables et réunir les renseignements nécessaires par entretien et observation directs. Outre son coût élevé, le procédé a le défaut d'omettre les activités non manifestes.
- b) Dossiers officiels. Les dossiers tenus par les administrations aux fins fiscales et pour la mise en oeuvre des programmes d'assurance-chômage, sécurité sociale ou autres peuvent aider à établir un répertoire. Ces dossiers, toutefois, doivent faire l'objet d'un examen attentif pour déterminer s'ils sont complets, exacts et propres aux fins que se propose le recensement, car ils ne sont pas destinés essentiellement à satisfaire les besoins des enquêtes économiques. Ils peuvent ne contenir aucune indication de taille et, ainsi, être impropres au choix d'échantillons ou à l'application d'une limite de taille pré-déterminée. Ils peuvent se rapporter aux entreprises et non aux établissements, ce qui oblige à une enquête préliminaire (souvent appelée tournée préalable) pour réunir les renseignements nécessaires sur chaque établissement. Les adresses peuvent être celles pour la correspondance et différer de la situation géographique. Il se peut que ces dossiers ne soient pas mis à jour et comprennent ainsi des entreprises qui ont changé de propriétaires ou qui ont fermé et en émettent d'autre récemment ouvertes. La répartition par genre d'activité (à supposer que ce renseignement figure dans le dossier) a pu s'opérer en fonction de règles de classement différentes de celles servant dans l'enquête. L'un des principaux avantages des dossiers administratifs est de comprendre des listes d'activités non manifestes et aussi d'autres qui pourraient échapper aux enquêteurs parce qu'elles sont difficiles à localiser ou qu'ils pourraient par erreur omettre d'inscrire parce qu'ils auraient mal compris leurs instructions.
- c) Répertoires établis par des associations ou des publications professionnelles. Etant donné que ces répertoires sont conçus à des fins déterminées, ils présentent certains défauts dont

doivent être avertis ceux qui les utilisent. Outre poser les problèmes qui découlent de l'utilisation de dossiers administratifs, les registres des associations professionnelles auront tendance à se borner à leurs membres. Pourtant, il arrive souvent qu'y figurent aussi des entreprises servant la branche considérée. Ainsi, une liste consacrée aux salons de beauté peut comprendre des grossistes et fabricants de produits de beauté. De même, étant donné que les publications professionnelles sont portées à s'assurer une clientèle aussi nombreuse que possible, leurs listes comprendront maintes entreprises qui n'ont que des relations marginales avec la profession.

- d) Autres sources éventuelles. Elles comprennent les annuaires téléphoniques, listes spéciales dressées par les services du téléphone, listes dressées par des sociétés privées spécialisées dans cette tâche, chambres de commerce, etc. Chaque genre présente ses caractères propres qu'il y a lieu d'étudier soigneusement avant de décider de la façon de l'utiliser.

155. L'un des principaux problèmes que pose l'établissement d'un répertoire en vue d'une enquête sur la distribution et les services tient au fait que les entreprises de cette branche ont souvent tendance à opérer en plusieurs emplacements différents. En outre, certaines se livrent à plus d'un genre d'activité. Pour donner toute son utilité au répertoire, il est nécessaire de chercher à reconnaître ces entreprises et à les contacter avant le début de l'enquête, pour pouvoir établir, pour chaque emplacement, une fiche séparée mentionnant le lieu, le classement par genre d'activité et un indicateur de taille. Pareille prise de contact ou tournée préalable peut s'opérer sur un choix d'entreprises connues pour présenter ces caractères (quand on dispose de ce renseignement).

156. En général, le répertoire s'élabore à raison d'une fiche par établissement. Pour les entreprises à succursales multiples, il faut aussi une fiche pour le siège central, avec renvoi pour chaque établissement. Le fichier peut consister en cartes ordinaires, un par établissement ou entreprise, ou bien on peut préparer un jeu de cartes perforées qui permet de passer sur ruban magnétique dans les pays disposant d'un système d'informatique. L'emploi d'une fiche par exploitation assure le maximum de souplesse et permet la suppression des fiches concernant les établissements ou les entreprises qui cessent leur activité, l'insertion de nouvelles fiches pour ceux et celles qui s'ouvrent, l'exécution de toutes modifications et le classement par ordre alphabétique ou numérique, par genre d'activité et par secteur. Pour les pays qui disposent d'un petit répertoire, il paraît indiqué de tenir un fichier de cartes ordinaires; l'emploi de cartes perforées assure un tri rapide et une grande souplesse. L'ordinateur est encore plus rapide et plus souple, mais il ne doit s'utiliser que si le volume et la complexité des tâches le justifient.

157. Toutes les sources servant à constituer un répertoire posent un même problème : celui de l'intervalle entre le temps où s'obtient le renseignement

et celui où s'utilise le répertoire. Il s'ensuit que les renseignements sont souvent incomplets et ne tiennent pas compte des changements intervenus entre-temps. Si le répertoire doit servir à des enquêtes successives, il faut prévoir quelque moyen de le compléter par l'inscription des entreprises nouvelles et de le corriger pour tenir compte des fermetures et des changements de propriétaires. Il y a donc lieu d'insérer dans le système un procédé permettant d'apporter ces changements. On veillera à ce propos à éviter toute double inscription des entreprises qui n'ont changé que de raison sociale ou qui en ont absorbé une autre.

158. Comme on l'a dit, la meilleure source pour l'établissement d'un répertoire complet s'obtient par exécution d'un recensement général. Dans ce recensement, on consacre beaucoup d'effort à reconnaître et classer chaque unité statistique et chaque unité déclarante, et on réunit des renseignements sur la taille, le lieu et autres particularités. Un répertoire ainsi établi peut servir à dresser des listes de déclarants qui serviront aux enquêtes annuelles ou plus fréquentes encore effectuées par tirage d'échantillons stratifiés. Ce mode de sondage peut se conjuguer avec un sondage aréolaire. Pareil répertoire devra probablement se limiter aux principaux entreprises et établissements, les autres faisant l'objet du sondage aréolaire; ce sondage peut alors servir à représenter les petits établissements aussi bien que les nouveaux. La liste des grandes entreprises à inclure dans l'enquête ne doit pas se borner à celles qui rentrent dans les segments sondés, et les données qui les concernent s'obtiendront directement par correspondance où que soient situés les établissements. La limite de taille choisie pour définir les grands établissements dépendra du compromis optimum entre le coût et les problèmes que posent respectivement le dénombrement sur place et le dénombrement par correspondance. On dressera la liste des grands établissements (et des entreprises à succursales multiples) classés par secteurs, qui sera transmise aux services extérieurs et indiquera aux enquêteurs les établissements à ne pas dénombrer dans les secteurs choisis pour le sondage aréolaire.

159. Etant donné qu'un répertoire énumère un ensemble d'entreprises en activité lors de sa préparation et qu'il ne saurait valoir mieux que la source d'où il provient et que le soin apporté à l'établir, il y a constamment lieu soit de le mettre à jour, soit de le compléter pour y inscrire les entreprises et établissements nouveaux. C'est à quoi sert le sondage aréolaire qui l'accompagne. Si on dispose d'autres sources telles que dossiers fiscaux et administratifs, on peut y puiser pour le tenir à jour. Toutefois, le dépouillement et la transcription de tels dossiers en temps opportun sont extrêmement difficiles.

160. Le répertoire une fois établi, certains pays élaborent un programme de réunion de données à partir d'un ensemble de grandes entreprises qui y figurent. Quand on procède ainsi, la première condition est d'élaborer un système où l'évolution, d'une période à l'autre, s'évalue à partir de déclarations faites par les mêmes unités. Ce programme présente des défauts manifestes : ce qui se produit pour les grandes entreprises peut n'être pas représentatif de ce qui se produit dans l'ensemble de l'économie et, pour les entreprises non déclarantes, l'évolution peut être

différente. Aussi, bien que ce genre de programme soit le moins coûteux dans son application, les données qu'il fournit risquent d'être trompeuses.

161. Comme on l'a dit, la meilleure source se trouve dans un recensement initial d'ensemble comprenant l'établissement d'une liste sur les lieux. Dans les recensements ultérieurs, la liste des établissements identifiés lors du premier sert de base au nouveau répertoire. Toutefois, dans un recensement, un tel répertoire ne s'emploie pas tout à fait comme dans une enquête annuelle ou plus fréquente. Il sert de moyen de contrôle, essentiellement pour s'assurer que les grands établissements et entreprises n'ont pas été omis, n'ont pas été classés différemment, ou n'ont pas été attribués à un secteur différent. Avant dénombrement, les données concernant les grands établissements et provenant du recensement précédent peuvent être classées par secteurs et communiquées aux bureaux extérieurs compétent pour leur servir à contrôler le dénombrement en cours. Les enquêteurs seront alors tenus de rendre compte des établissements non localisés. Une fois le dénombrement achevé, les formulaires pourront être classés selon le lieu, le genre et le volume des activités, et on collationnera les données provenant des deux recensements successifs. Les différences et omissions importantes seront élucidées. On veillera toutefois à s'assurer que le défaut de concordance ne tient pas à un changement de nom ou de propriétaire.

162. Un répertoire établi en vue d'un dénombrement par correspondance joue un rôle différent. Il doit alors être complet en soi, c'est-à-dire qu'il doit comprendre une fiche plus ou moins détaillée pour chaque unité dont on utilisera les données. La première condition est donc que les sources soient complètes et qu'elles contiennent les données nécessaires au recensement ou permettent de prendre contact avec l'entreprise pour les obtenir. Etant donné que le recensement se borne à certains genres d'activité, la source doit indiquer le genre d'activité aussi bien que le nom et l'adresse. Il serait aussi très utile d'avoir des renseignements sur la taille. Le mieux serait que le répertoire comprenne une fiche séparée pour chaque établissement d'une entreprise à succursales multiples. Quant un répertoire réunit tous les éléments ci-dessus, il peut servir non seulement à contrôler le dénombrement, mais aussi à attribuer des données par interpolation dans tous les cas d'abstention et de négligence. A partir d'un tel répertoire, on peut mettre en tableaux les comptages d'établissements par genre d'activité, lieu et classification par taille. Sans être aussi précis que ceux qu'on peut tirer d'un formulaire dûment rempli, ces comptages peuvent se révéler utiles à diverses fins. Lors du dénombrement, il sera possible, en effet, de préparer une cédule pour les abstentionnistes et contrevenants, en attribuant à leurs établissements les mêmes caractéristiques que celles obtenues en tabulant les données fournies par ceux de même taille et même genre d'activité, sis dans le même secteur général.

163. Quand, pour une même entreprise, différents fichiers officiels utilisent un système commun d'identification, il est également possible d'en transcrire les données effectives (d'ailleurs d'ordinaire très limitées) sur le répertoire. Par là, on peut libérer toute une catégorie de petites entreprises de l'obligation de déclarer. Pour utiliser ainsi

ces fichiers administratifs, il convient d'étudier plusieurs éléments pour s'assurer de la validité du procédé. Un, il doit être possible de déterminer l'unité statistique ou déclarante; deux, le classement par genre d'activité doit se fonder sur le même système; trois, les données à transcrire doivent être définies de la même façon qu'elles le seront dans l'enquête;; quatre, le fichier administratif doit être à jour.

Collecte et contrôle des déclarations

164. Dans une enquête qui exige une visite aux établissements qu'elle concerne, les renseignements réunis pendant qu'on en dresse la liste peuvent être complétés de façon à fondre listage et collecte. Comme on l'a dit, étant donné que la distribution et les services se caractérisent par un grand nombre de petits établissements, un léger surcroît d'informations obtenues en dressant la liste peut servir à faire de celle-ci le moyen de réunir les données concernant ces petits établissements pour lesquels les renseignements essentiels suffisent. Pour tous les autres, l'enquêteur pourra laisser le questionnaire approprié, soit en fixant la date où il reviendra le ramasser, soit en invitant à l'expédier au bureau local dont il dépend. Ayant achevé son dénombrement selon la liste, l'enquêteur remet son registre à ce bureau, qui le collationne alors avec les déclarations qu'il a reçues soit des mains de l'enquêteur, soit par la poste.

165. En dressant la liste, il faut obtenir pour chaque établissement un indicateur de taille. Il servira à l'attribution de données par interpolation aux abstentionnistes. Il indiquera aussi utilement la peine qui se justifie pour tenter d'amener les négligents à répondre.

166. S'il a affaire à des établissements appartenant à des entreprises à succursales multiples, l'enquêteur doit obtenir le nom et l'adresse du siège qui fournira les données. Pour ces établissements, des fiches de renvoi seront préparées et expédiées au service statistique central qui, en les conférant avec les fiches analogues provenant du reste du pays, pourra contrôler les déclarations de ces entreprises.

167. Si le listage a été convenablement conduit et contrôlé, les établissements inscrits comme rentrant dans l'enquête forment l'univers statistique à étudier. Les opérations ultérieures devront s'appliquer à cet univers si l'on veut que l'enquête se déroule régulièrement.

168. Pour des raisons pratiques, il est souhaitable de déceler et de corriger les erreurs entachant les déclarations dès que possible après que le questionnaire ou le registre a été rempli. Si l'entrevue ou l'entretien ultérieurs tarde, il se peut que la personne qui a fourni les premiers renseignements ne travaille plus dans l'entreprise, que celle-ci ait cessé son activité ou qu'elle n'ait pas conservé ses livres. En revanche, il est d'ordinaire ni possible ni souhaitable d'opérer un collationnement complet dans les services locaux. Leur personnel doit toutefois être capable d'opérer un contrôle comprenant une certaine vérification des données. Il peut consister à s'assurer que le questionnaire est convenablement attribué, qu'aucun chiffre ne manque et autres contrôles

simples. Un contrôle un peu plus complexe, quoique essentiel, consiste à additionner les dépenses déclarées (traitements et salaires, matériaux, etc.) pour s'assurer qu'elles ne dépassent pas la valeur déclarée des ventes ou services. Sinon, le formulaire doit être rendu à l'enquêteur ou donné à un autre pour vérification ou correction.

Pareil examen permet aussi de contrôler quelque peu les résultats obtenus par les divers enquêteurs. Quand les questionnaires ou registres se révèlent défectueux, l'enquêteur responsable doit être invité à les corriger, si le temps le permet. Si leur nombre était excessif, il signifierait que l'enquêteur est incompetent et devrait être licencié.

169. Le répertoire complet et à jour des établissements et autres unités déclarantes étant établi, il peut servir à choisir les unités à inclure dans l'enquête, qu'elle se fasse par dénombrement intégral ou par sondage. Ces unités seront choisies selon leur genre d'activité et dénombrées au moyen d'un questionnaire particulier pour chaque groupe d'activité. Si l'enquête doit se faire par dénombrement postal, l'adresse indiquée dans le répertoire y servira et l'enquêté sera invité à retourner dans un délai donné le questionnaire dûment rempli au service statistique central, qui l'enregistrera à réception. Périodiquement, il sera nécessaire de relancer les entreprises négligentes. Une exécution prompte et efficace exige la mise au point d'un ensemble serré de contrôles, qui seront strictement appliqués. Sinon, on risque fort des omissions et doubles emplois.

170. Le questionnaire approprié selon le genre d'activité, envoyé à l'entreprise, doit s'accompagner d'autres documents. Ils comprennent une enveloppe marquée "franchise postale" portant l'adresse du service auquel l'enquêté doit retourner le questionnaire dûment rempli, un double de ce questionnaire qu'il pourra conserver pour ses propres dossiers et toutes instructions particulières dont il a besoin pour le remplir et le retourner; plus l'affirmation que les déclarations seront traitées de façon confidentielle et l'énoncé des peines applicables en cas de non-déclaration ou de fausse déclaration. Des enveloppes portant des marques distinctives attireront l'attention de l'enquêté sur l'importance de leur contenu. L'emploi d'enveloppes ajourées laissant voir l'adresse portée sur le questionnaire dispense de la répéter sur l'enveloppe. La date à laquelle retourner les déclarations doit être clairement indiquée. Il importe de fixer un délai qui laisse un temps suffisant pour remplir le questionnaire, mais qui ne soit pas tellement long que l'enquêté risque de le mettre de côté et de l'oublier. Généralement, on peut considérer qu'un mois est un délai convenable. Il ne sera possible de prescrire un jour précis que si les questionnaires sont prêts à être expédiés par les enquêtés à peu près à la même date. Si ce n'est pas le cas, on leur demandera de les renvoyer dans les 30 jours après réception. De même, les pays où la poste laisse à désirer peuvent éprouver des difficultés à prescrire un jour précis et pourront devoir demander de remplir et de retourner le questionnaire dans les 30 jours de sa réception. En pareil cas, il devient très difficile de relancer les négligents.

171. Pour s'assurer qu'aucun établissement figurant sur la liste n'est négligé, de façon que les tableaux définitifs représentent intégralement l'univers statistique considéré, il faut établir un système serré de contrôle. A cet effet, on commence par fichier, sur cartes ordinaires, cartes perforées ou rubans magnétiques, chaque unité à qui un questionnaire a été expédié. L'attribution d'un numéro matricule unique à chacune y aidera, car la raison sociale risque de changer, d'être confondue avec une homonyme, d'être mal orthographiée, de suivre un changement de propriétaire, etc. Ce numéro matricule peut aussi servir au classement final des déclarations et simplifiera leur recherche si elles sont nécessaires au cours de la révision des tableaux. Deux genres de listages se révéleront très utiles durant le dépouillement : l'un par ordre numérique et l'autre par ordre alphabétique. L'emploi d'un fichier de cartes procure une grande souplesse, malheureusement les cartes risquent de se perdre ou d'être mal classées. Les listes ne présentent pas ce défaut, encore qu'on doive veiller à les assembler de préférence dans des classeurs, pour les protéger.

172. A mesure qu'on reçoit les déclarations des enquêtés, on les enregistre pour pouvoir éplucher le dossier tout entier lorsqu'on jugera nécessaire de relancer les négligents. L'effort de relance requis pour obtenir une proportion convenable de réponses est grand, même dans les pays avancés. Si l'on a fixé un délai d'un mois après l'expédition des questionnaires, on ne peut pas compter que plus de la moitié soient prêts en temps voulu. Comme il n'est pas possible de présenter les données sans y inclure les principales entreprises, il est bon de fixer à leur intention des règles spéciales. On doit leur rendre visite avant l'enquête pour régler tout cas particulier éventuel. En cours d'enquête, on s'efforcera de suivre attentivement leurs communications pour s'assurer qu'elles sont reçues suffisamment à l'avance et que leurs déclarations sont acceptables. Préparer ces déclarations peut leur prendre plus de temps et le délai peut devoir être prolongé pour ces entreprises. De même, on pensera à prolonger convenablement le délai pour celles qui ont des problèmes particuliers. Quand un cabinet comptable est chargé d'établir un nombre important de déclarations, il doit également disposer de plus de temps.

173. L'envoi des formulaires par la poste suscitera une abondante correspondance, car les enquêtés poseront souvent des questions dont ils attendront la réponse pour remplir le questionnaire. Un procédé approprié doit être élaboré pour trier ces lettres en vue d'y répondre promptement.

174. En choisissant le système qui sera appliqué à la collecte des données, il importe de se souvenir que l'un des principaux objets de pareilles enquêtes est d'obtenir des séries chronologiques. Les mouvements cycliques à long terme ressortent de l'étude des recensements; les canevans courants apparaissent dans les enquêtes qui ont lieu plusieurs fois par an. Dans les deux cas, toutefois, l'authenticité de ces mouvements peut se trouver infirmée par une modification de la méthode appliquée. En concevant l'enquête, il faut adopter un système qui puisse être conservé d'une exécution à l'autre. Si, par exemple, on décide, pour établir le répertoire, d'utiliser les dossiers d'une administration donnée, on devra être assuré

que cette administration continuera à les tenir de façon acceptable dans un avenir prévisible. De même, en décidant de l'ampleur et de la nature de l'échantillon, on devra avoir quelque assurance que l'opération pourra se poursuivre. Toute "amélioration" doit être examinée pour voir si elle n'entraîne pas de discontinuité dans les données. Toutefois, s'il est jugé nécessaire d'adopter un procédé qui entraîne une telle discontinuité, on s'efforcera par tous les moyens d'établir un "passage" entre l'ancienne série et la nouvelle, de façon que les utilisateurs soient avertis de l'effet du changement et puissent apporter les correctifs nécessaires.

Recours au sondage^{6/}

175. Qu'il s'agisse d'un recensement complet embrassant un grand nombre de données, ou d'enquêtes soit annuelles soit plus fréquentes encore destinées à fournir le plus vite possible des données moins nombreuses, le responsable cherchera des modalités qui procurent les données requises au moindre coût. Le sondage étant l'une des méthodes qui réduisent les coûts, il y a intérêt à considérer chaque programme pour déterminer dans quelle mesure et de quelle façon on peut y recourir. Pour en décider, le responsable doit peser les avantages offerts par le sondage, tels que réduction du nombre des contacts à prendre, des questionnaires à dépouiller et du personnel nécessaire, avec une publication plus prompte des données, contre les inconvénients, tels que difficultés inhérentes au contrôle des échantillons, qualification requise du personnel et renoncement à certaines données. Dans les enquêtes sur la distribution et les services, l'ampleur et la nature des échantillons à utiliser dépendent de l'hétérogénéité des genres d'activités à traiter. Par exemple, un sondage qui conviendrait pour mesurer l'activité du commerce de détail serait probablement impropre à mesurer celle d'une agence de publicité. Un plan de sondage qui serait acceptable pour un recensement général pourrait mal fonctionner pour une enquête mensuelle.

176. Quand la méthode d'enquête comprend un listage sur les lieux, on peut fixer une limite de taille au-dessus de laquelle tous les établissements sont dénombrés et au-dessous de laquelle on sonde tout nième établissement. Cette limite peut être tracée à partir de lignes préalablement désignées dans un registre ou choisie après dépouillement de ces registres au service statistique central. Multipliés par un coefficient approprié, les établissements sondés représentent alors l'ensemble de ceux qui se trouvent au-dessous de la limite. De telles données seront évidemment entachées d'une erreur de sondage. Au lieu de choisir pour le sonder tout nième établissement au-dessous de la limite, une variante consiste à réunir à son sujet les renseignements complets et de ne remettre aux autres qu'un questionnaire très simplifié qui se borne aux données fondamentales. Tous les établissements situés au-dessus de la limite seraient pour leur part tenus de fournir des renseignements complets.

^{6/} On trouvera à l'Annexe I un exposé détaillé des techniques de sondage.

177. Quand un dénombrement postal s'opère au moyen d'un répertoire qui contient des renseignements sur la taille et le genre d'activité, il est possible de choisir un échantillon stratifié fondé sur ces caractères. Dans le plan d'enquête, toutefois, il importe d'inclure tous les grands établissements. Etant donné le nombre élevé des petits, il est possible d'adopter un plan qui réduise considérablement le nombre des unités sondées, mais un tel échantillon réduira d'autant la possibilité de publier des renseignements par secteur et genre d'activité.

178. Les données visées par les enquêtes annuelles et plus fréquentes encore sont d'ordinaire très limitées. En outre, il importe de les diffuser promptement, surtout dans les enquêtes qui ont lieu plusieurs fois par an, puisqu'il s'agit de réunir des renseignements concernant les tendances courantes qui serviront de guide à des mesures immédiates. Il en va particulièrement ainsi pour des rubriques telles que les ventes au détail. En conséquence, l'enquête doit être d'un plan aussi simple que possible et se prêter à un dépouillement rapide. Pour qu'elle soit valable, elle doit s'opérer par sondage probabiliste dont l'échantillon est choisi avec une probabilité connue, de façon qu'on puisse calculer sa variabilité. Le plan le plus commun consiste à dresser une liste des grandes entreprises qui seront sondées au cours de chaque période d'enquête et à partir de laquelle on pourra élaborer un système de choix de l'échantillon de toutes les autres. Ce sera soit un échantillon de superficies (aréolaire) ou, si on dispose d'une liste complète et à jour des entreprises, un échantillon tiré de cette liste. Comme il est quasi impossible de tenir cette liste à jour, il faudra utiliser en complément un échantillon aréolaire qui pourra être bien plus petit que si cette liste manquait. Bien que le rôle essentiel de l'échantillon aréolaire soit de représenter les petits établissements, il sert encore à une fin importante : trouver et dénombrer les établissements nouveaux, quelle que soit leur taille, apparus après l'inclusion des "grands" établissements dans l'échantillon.

179. Pour quelques genres d'activités, un sondage aréolaire est sans effet et entraîne des variances si fortes qu'elles limitent ou même proscrivent son emploi. La cause en est que l'objet principal de l'enquête est d'élaborer des données sur le mouvement du commerce de détail, car c'est là l'un des indicateurs économiques les plus sensibles. En conséquence, le choix des segments de l'échantillon est centré sur les agglomérations qui correspondent exactement à l'exercice de ce commerce. En revanche, des activités telles que celles des hôtels saisonniers, des stations de stockage de produits pétroliers, des marchands de ferraille en gros, des producteurs de films cinématographiques et des stations d'émission de radio, également comprises dans l'enquête, ne sont pas ordinairement situées dans des agglomérations : par suite, la partie des données réunies dans les segments de l'échantillon présentera des variances si fortes qu'elle les rendra sans valeur. Pour ces activités, il est donc nécessaire de se fonder essentiellement sur un répertoire assez complet qu'on puisse tenir à jour. Si ce n'est pas possible, ces activités seront exclues des enquêtes courantes.

180. En théorie, étant donné que par sa moindre taille, un échantillon réduit le nombre des déclarations à réunir et à dépouiller, il présente un certain nombre d'avantages. Bien des enquêtes sont ainsi déchargées du

fardeau de la déclaration, et la tâche du service statistique se trouve allégée d'autant. Il peut donc exercer sur la collecte et le dépouillement des déclarations un contrôle plus serré qui aboutit à améliorer la qualité des statistiques obtenues. En outre, leur présentation gagnera en ponctualité et il deviendra possible de choisir la taille de l'échantillon en fonction des crédits disponibles. Quand ces crédits ne suffisent pas à un recensement complet, le sondage peut être la seule méthode praticable.

181. Ce sondage doit être un sondage probabiliste qui donne des résultats d'une qualité mesurable. Réaliser un tel sondage requiert des démarches supplémentaires qu'il convient d'exécuter convenablement. Elles entraînent quelque dépense et créent des risques d'erreur qui ne se présentent pas avec un dénombrement complet. A moins que le sondage ne soit conçu pour s'ajuster aux conditions d'exécution en vigueur et qu'il fasse l'objet d'un contrôle satisfaisant, on peut en attendre plus d'inconvénients que d'avantages. Un plan de sondage perfectionné et théoriquement opportun, qui exige plus de personnel qualifié, de formulaires ou de matériel qu'on n'en dispose serait vain ou pire encore. Le sondage présente un autre inconvénient : il réduit inévitablement le volume des détails fiables qu'on peut publier, car le petit nombre d'observations dont on dispose pour chaque rubrique accroît la variance et, dans de nombreux cas, la porte à un degré tel qu'il fausse complètement les statistiques obtenues. De plus, bien qu'on puisse mieux contrôler la collecte, le collationnement et la présentation grâce au moindre nombre de déclarations à exploiter, les erreurs dans l'exécution de ces tâches ne sauraient être complètement éliminées.

VIII. TRAITEMENT DES DONNEES

Observations générales

182. Après retour des questionnaires, il y a lieu de les examiner pour s'assurer qu'ils contiennent tous les renseignements nécessaires et qu'ils sont en état d'être traités. Cet examen s'assurera notamment que les données requises ont été fournies, que les chiffres ont été arrondis et les rubriques codées comme il convient par le personnel, que les inscriptions sont lisibles, qu'il est donné suite aux observations des enquêtés et qu'on a procédé aux collationnements pour déceler les données erronées. Cet examen fera apparaître un grand nombre de problèmes qu'il y a lieu de résoudre avant que les déclarations soient en état d'être dépouillées. On devra prendre contact avec un pourcentage élevé d'enquêtés pour vérifier ou corriger les données. Cette tâche prenant beaucoup de temps, elle devra commencer tôt et être menée vivement pour ne pas retarder les opérations ultérieures.

183. La tâche la plus difficile est de déceler les erreurs dans les déclarations. Ces erreurs peuvent tenir à diverses causes : l'enquêté ne comprend pas les notions inhérentes à une déclaration exacte; le déclarant n'est pas soigneux; l'établissement ne disposait pas des données requises; les données fournies sur l'emploi et les salaires, par exemple, se rapportent à des périodes différentes, etc. Ces problèmes doivent être constatés et résolus avant l'exploitation des données. Le volume de travail à exécuter par le personnel de bureau et par les statisticiens variera selon la qualité de ce personnel et le matériel disponible. Grâce aux ordinateurs, des calculs complexes destinés à déceler les erreurs peuvent s'exécuter à vive cadence. La tâche du personnel de bureau et des statisticiens sera d'autant plus lourde que le matériel est moins perfectionné.

184. Traiter de même toutes les déficiences des réponses serait du gaspillage et il convient d'établir quelques normes et critères de collationnement. En général, l'effort consacré à déceler et à corriger les erreurs et omissions doit être proportionné à leur effet probable sur les données publiées. Les chiffres offrent une certaine sécurité : si les erreurs sont de caractère aléatoire, elles auront tendance à s'annuler, pourvu que chacune soit faible. Il en découle qu'une grande précision et des normes strictes ne s'imposent pas à propos des petits établissements, qui caractérisent une bonne partie du commerce de détail et des services. Le degré d'attention portée à chaque questionnaire pendant son collationnement et son codage devra varier en raison directe de la taille des établissements représentés. Il s'ensuit que les déclarations des grands établissements seront normalement collationnées avec plus de soin que celles des petits, sauf quand l'enquête sur les petits se fait par sondage.

Simplees opérations de bureau

185. La répartition des tâches, la tenue des registres et les rapports d'avancement seront simplifiés si les questionnaires sont classés par

genre et réunis en liasses de taille appropriée. D'ordinaire, une liasse n'en comprendra pas plus de 100; le nombre exact dépendra du volume de travail à exécuter mais ne devra pas être tel qu'il bloque une quelconque liasse pendant un temps prolongé. Les liasses seront numérotées et, pour en suivre l'avancement, on tiendra un registre central indiquant pour chacune la date prescrite, le nom du commis qui s'en occupe, le nombre de questionnaires obligeant à reprendre contact avec l'enquêté, le nombre de questionnaires transmis pour analyse statistique et la date d'exécution. Une liasse ne comprendra d'ordinaire qu'un seul type de questionnaire. Il peut être bon de classer les questionnaires d'autres façons, par exemple en séparant ceux qui proviennent des petits établissements et ceux qui proviennent des grands, si chaque catégorie a ses propres règles de collationnement. De même, les rapports provenant des entreprises à succursales multiples pourraient former une série séparée de liasses.

186. L'une des principales décisions à prendre à ce stade consiste à fixer jusqu'à quel point la correction et le collationnement seront divisés en opérations spécialisées : chaque commis peut soit se borner à répéter quelques tâches très simples qui s'apprennent vite, soit apprendre et exécuter toutes les opérations nécessaires pour préparer le questionnaire en vue de la mise sur cartes perforées. La première solution présente l'inconvénient d'obliger à un contrôle assez complexe et à de fréquentes manipulations des formulaires. La seconde demande du temps pour former un personnel qualifié pour exécuter les opérations plus complexes et plus difficiles. La décision sera affaire de discernement et dépendra du personnel disponible. Il peut être indiqué d'adopter un compromis.

187. Ces simples tâches de bureau se prêtent bien à l'application d'un contrôle qualitatif destiné à signaler promptement les commis qui n'appliquent pas convenablement les instructions. Ce contrôle devrait aussi être conçu de façon à mesurer leur rendement en même temps que la qualité. Les commis seraient avertis immédiatement de leurs erreurs et ceux qui demeurent au-dessous des normes acceptables seraient licenciés. Souvent, ce genre de contrôle peut aussi révéler une erreur commise dans les instructions, et risquant d'égarer tout le personnel chargé du collationnement.

188. Il incombe au personnel chargé de la planification de préparer des spécifications qui serviront au personnel d'exploitation à rédiger pour les commis des instructions sur les critères permettant de déterminer si la déclaration est acceptable ou défectueuse. Ce personnel doit aussi fournir la documentation de référence qui permet d'interpoler les renseignements manquants et les systèmes de codage à appliquer aux données. Comme toujours, le même personnel doit examiner les résultats obtenus. Il doit s'arranger pour sonder certaines des premières liasses aussitôt après leur mise en état et apporter dès que possible toutes corrections de méthode qui sont nécessaires. Il est souhaitable aussi qu'il passe périodiquement dans les bureaux pour voir si de nouveaux problèmes apparaissent.

189. Un service de correspondance devrait être assuré au titre des opérations de correction et de collationnement. On constatera en effet

que bien des réponses nécessaires pour rendre le questionnaire complet ou acceptable peuvent être demandées et obtenues au moyen d'une circulaire. Le commis pourra indiquer à cet effet la circulaire à envoyer (parfois, on peut en rédiger une seule qui comprenne plusieurs points) et enverra un bordereau portant des indications au service de correspondance qui expédiera la lettre et établira des contrôles pour traiter les réponses.

190. Dans certains cas, on ne pourra prescrire de procédé simple pour ces opérations : il se présentera certaines situations exigeant une méthode plus complexe. On constituera alors un second groupe de commis plus qualifiés et spécialement formés pour s'en occuper. Les problèmes présentant un caractère unique et insolubles par l'un ou l'autre groupe doivent être transmis aux statisticiens de profession. Ceux-ci auront le choix d'élaborer eux-mêmes la solution ou de reprendre contact avec l'enquêté pour résoudre le problème.

191. Parmi les règles qui présentent de l'intérêt et servent en maints pays figurent les suivantes :

- a) Les mentions portées sur le questionnaire par les commis devraient être faites au crayon ou stylo à bille de couleur déterminée. Cette couleur doit contraster avec celle des inscriptions portées par les enquêtés ou les autres personnes qui manipulent le formulaire.
- b) Les mentions portées d'après la réponse d'un enquêté à une demande de précisions doivent aussi être de couleur distincte.
- c) Pour corriger la réponse d'un enquêté, il ne faut ni la gommer ni l'effacer. Le commis la barrera de façon qu'elle reste lisible et portera au-dessus le chiffre corrigé. Si la place manque, ce chiffre sera porté ailleurs, une flèche indiquant la place qu'il doit occuper.
- d) Sur le formulaire doit être prévu un espace pour les initiales du personnel (de collationnement ou de codage, ou encore statisticiens) qui s'est occupé des formulaires, avec la date.

Préparation des questionnaires en vue de les tabuler

Vérification et contrôle

192. La première démarche consiste à déterminer si le questionnaire concerne un établissement qui rentre dans le secteur de la distribution et des services. Cette opération s'impose tout particulièrement quand le questionnaire lui a été envoyé par la poste et qu'on doute de son genre d'activité. S'il est déclaré hors secteur, aucune suite n'est donnée. Etant donné qu'une telle décision repose sur un examen de son genre d'activité, il pourrait sembler logique à ce point de lui attribuer la cote correspondante. Toutefois, les règles de ce codage risquent d'être d'application très complexe, et il serait bon de confier cette opération à un

groupe de commis spécialement formés à cet effet. Pour attribuer de telles cotes, il sera nécessaire de se fonder sur des renseignements tels que les ventes par catégorie de client, le volume des recettes provenant des services fournis par opposition à celles provenant des ventes de marchandises, les ventes par catégorie de marchandise, les services fournis (par genre), le genre d'activité que s'attribue l'établissement et une cote historique (si elle existe). Les modes d'utilisation de ces renseignements pour classer selon le genre d'activité seront fournis par le personnel chargé de la préparation de l'enquête, qui se tiendra aussi à disposition pour résoudre les nombreux problèmes qui pourraient se poser.

193. Plusieurs rubriques doivent figurer pour chaque établissement inclus dans le champ de l'enquête. Elles comprennent d'ordinaire le volume total des ventes, l'effectif global occupé et le montant annuel des salaires versés. (Dans les enquêtes qui ont lieu plusieurs fois par an, les données nécessaires peuvent se borner au volume des ventes et comprendre des inventaires dans le cas des commerces de gros et de détail). Si les grands établissements n'ont pas fourni les données requises, il sera nécessaire de les relancer. Pour les petits, en revanche, il peut être plus indiqué d'interpoler. Un jeu de tableaux sera préparé pour servir de base à chaque interpolation.

194. Une situation particulièrement préoccupante se produit quand on demande une réponse à une question sur une activité que ne pratiquent pas tous les enquêtés. L'absence de réponse présente alors un caractère d'ambiguïté, car elle peut signifier soit que l'enquêté ne se livre pas à cette activité, soit qu'il ne s'est pas soucié de fournir le renseignement. L'exploitation des réponses peut alors se trouver faussée si la situation de ceux qui ont effectivement répondu diffère de celle des autres, surtout si l'omission signifie l'absence d'activité. S'il importe particulièrement de distinguer entre ceux à qui la question ne s'applique pas et ceux qui ont négligé d'y répondre, elle devra appeler une réponse par "oui" ou par "non", comme suit :

Avez-vous tiré un revenu de ... pendant l'année du recensement .
 Oui Non

Si la réponse est "oui", combien vous a rapporté cette activité pendant l'année du recensement ?

Si le "oui" est coché sans mention du montant, il est clair que l'enquêté a négligé de l'indiquer.

Collationnement et codage

195. Parce qu'elles se révèlent souvent fausses, les données déclarées doivent faire l'objet d'un collationnement critique pour déceler les cas douteux. Ce collationnement doit être double. Le premier vérifie la concordance interne des réponses fournies aux différentes rubriques par le même établissement, et le second s'assure que la réponse donnée à chaque rubrique est raisonnable en soi. Le contrôle de concordance interne consiste à vérifier si l'addition des éléments constitutifs redonne le total

déclaré, si les réponses données à la même question en différents points du formulaire sont les mêmes et si certains rapports numériques, comme celui des salaires à l'effectif ou encore aux ventes, rentrent dans des limites acceptables. Ces limites peuvent se déterminer à partir des moyennes tirées de données fournies par des établissements analogues de la même branche et du même secteur géographique. Si des enquêtes analogues ont déjà eu lieu, les données qui en proviennent peuvent aussi servir à cet effet. Toutefois, elles doivent alors être corrigées en fonction des changements connus qu'ont subis les prix et les salaires. Dans les enquêtes initiales, il pourra être nécessaire de se guider d'après des données provenant d'autres pays. Exemple de discordance interne : un établissement qui déclare être un magasin d'alimentation inscrit "zéro" à la rubrique des recettes provenant des ventes d'épicerie.

196. D'autres modes de contrôle s'imposent pour s'assurer que la réponse est plausible. On sera fondé à en douter pour des hôtels avec un taux d'occupation de "120 pour cent", des réparateurs de télévision sans recettes pour services, des salles de cinéma avec 10 sièges, des blanchisseries dans machines à laver, etc. En cours de collationnement, on s'assurera de l'efficacité des différents contrôles d'après le nombre des questionnaires rejetés comme fautifs. S'ils sont en majorité, les tolérances sont quasi certainement trop étroites. Si, en revanche, leur nombre est quasi nul, les tolérances sont probablement trop larges et les correctifs appropriés doivent être apportés.

197. Selon la mesure dans laquelle le questionnaire a été rédigé en vue du pré-codage, le commis devra examiner les renseignements qui y sont portés et attribuer une cote numérique qui pourra être transcrite sur carte perforée et servir à trier et tabuler. Même si les cotes sont imprimées d'avance sur le formulaire, il sera nécessaire d'examiner certains cas, comme celui de la réponse "manuscrite" à coder ou la présence de deux ou plusieurs mentions qui s'excluent mutuellement.

198. Les deux codages les plus difficiles sont celui du genre d'activité et du lieu. La détermination du premier oblige à demander et à obtenir par le questionnaire les renseignements appropriés et à appliquer des règles souvent complexes. Du codage du lieu, on peut penser qu'il se borne à attribuer une cote numérique à un toponyme, mais l'expérience enseigne qu'il présente plus de difficulté, surtout pour les villes. Les limites officielles sont difficiles à déterminer, et les enquêtés donnent souvent pour adresse un nom de ville, alors que leur établissement se trouve en fait en dehors de ces limites. Dans un dénombrement sur place, on peut surmonter la difficulté en demandant à l'enquêteur de déterminer l'emplacement; ce n'est pas possible dans une enquête postale. Le codage concerne encore la forme juridique de l'entreprise, le mode de propriété, les méthodes de vente, les catégories de marchandises, les diverses tailles, etc.

199. Les valeurs sont fréquemment déclarées dans la plus petite unité monétaire, mais il est souvent inopportun de perforer ou de tabuler de cette façon. En effet, perforer le montant ainsi indiqué prendrait du temps et exposerait à davantage d'erreurs; tabuler demanderait aussi plus

de temps et de place. Des problèmes analogues se posent quand les quantités sont indiquées jusqu'au dernier chiffre des unités. Une partie de la préparation du formulaire en vue de le tabuler consiste à réduire le nombre des chiffres en les arrondissant. Le personnel doit être soigneusement formé à cette opération : beaucoup semblent incapables d'assimiler les principes applicables. Certes, il serait plus facile de barrer les derniers chiffres que d'arrondir, mais il faut examiner les données considérées pour s'assurer que, par là, on n'influe pas de façon sensible sur les statistiques obtenues.

200. Dans le formulaire, on relèvera toutes notes ou observations portées par l'enquêté qui peuvent donner matière à correspondance. En dernier lieu, le commis devra revoir le formulaire pour s'assurer que les données à perforer sont lisibles et que le perforateur pourra sans difficulté les localiser rapidement.

201. Une enquête qui recourt largement à la poste pour distribuer et collecter les formulaires entraînera une masse de correspondance. Une bonne partie de celle-ci portera sur des questions courantes et pourra se faire au moyen de circulaires. Il n'en restera pas moins un gros travail pour régler les questions techniques qu'il y a lieu de résoudre pour que l'enquêté puisse compléter ses déclarations. Si cette tâche n'est pas confiée à un personnel convenablement formé pour traiter ces questions, les réponses peuvent facilement tarder pendant des semaines. L'achèvement de l'enquête en sera retardé et l'enquêté sera fréquemment découragé de compléter ses déclarations.

Collationnement et codage par ordinateur

202. Si la masse de travail à exécuter le justifie et si on dispose du genre approprié de personnel et de matériel, une bonne partie du collationnement peut s'opérer par ordinateur. Il n'en faudra pas moins que des commis revisent les déclarations pour s'assurer que les mentions essentielles y figurent, qu'elles ne présentent pas d'erreurs manifestes et que toute tâche qui ne peut être confiée à l'ordinateur est exécutée avant transmission pour mise sur cartes perforées. Après passage sur ruban, l'ordinateur peut exécuter bien des opérations, telles que calculs de coefficients, limites d'acceptation, interpolation des données et codage. Il peut aussi appliquer une série des critères d'acceptabilité et déceler les cas qui exigeront un contrôle ou, dans les cas appropriés, calculer le montant à inscrire à la place d'un autre manifestement erroné. En général, l'ordinateur peut s'acquitter de ces fonctions de façon plus rapide et plus précise que des commis.

203. On pense communément que l'informatique abrège le délai de publication. L'expérience ne justifie pas cette croyance. Selon une autre erreur commune, l'informatique entraîne automatiquement une forte réduction des effectifs. Tel n'est pas non plus nécessairement le cas. Avant de décider de collationner par ordinateur, il convient de considérer un certain nombre de points :

- a) La traduction des instructions, données par le personnel chargé de préparer l'enquête, en langage machine utilisé par le programmeur oppose un obstacle supplémentaire à la communication. A moins que ce personnel ne sache lire ce langage pour s'assurer que ses idées ont été convenablement interprétées, on s'expose à des erreurs. De toute façon, ce personnel devra avoir reçu une formation qui lui permette de suivre intelligemment le travail du programmeur.
- b) Bien des formulaires que l'ordinateur marque pour le réviseur doivent être corrigés. Or, il est coûteux de corriger des cartes perforées et des rubans magnétiques. Donc, si de très nombreux formulaires exigent finalement cette révision, la perforation de nouvelles cartes et ainsi de suite, le collationnement par ordinateur peut entraîner une dépense globale bien supérieure.
- c) L'ordinateur qui sera utilisé peut avoir une mémoire trop faible pour exécuter convenablement tous les contrôles de collationnement outre d'autres opérations, telles que coder et tabuler. Le programme doit être essayé en tant que de besoin pour déterminer l'effet de cette mémoire limitée.
- d) Un ordinateur de capacité suffisante offre des possibilités de contrôle bien supérieures à celles de commis. Le nombre de ces contrôles de concordance interne, par exemple, peut s'en trouver multiplié. Certains technocrates, enthousiasmés de cette extraordinaire capacité, ont imposé tant de contrôles de collationnement que presque chaque formulaire se trouve marqué pour révision. Comme dans le collationnement manuel, l'utilité de chaque contrôle doit être envisagée par rapport à son coût, comme par rapport à l'effet cumulatif global d'un grand nombre de contrôles.
- e) Par l'informatique, le collationnement intervient nécessairement plus tard. Comme on l'a déjà dit, il est souhaitable d'interroger les enquêtés aussitôt que possible après réception des déclarations. Dans un tel collationnement donc, il convient de tenir compte du retard imposé à cette reprise de contact.
- f) Une différence importante entre l'informatique et les méthodes qui l'ont précédée, à savoir dépouillement manuel, mécanique ou électromécanique, est que dans l'informatique toute fausse interprétation entraînant des erreurs itératives ou systématiques est souvent bien plus difficile à rectifier. Beaucoup d'onéreuses tabulations ont dû être mises au rebut parce que le programmeur n'avait pas bien compris ce que voulait l'organisateur responsable. Avec des anciennes méthodes, de telles erreurs se découvriraient et se rectifieraient d'ordinaire alors qu'il n'y avait encore que demi-mal. L'extrême rapidité de l'informatique fait, au contraire, que des projets entiers sont souvent achevés avant qu'on découvre les erreurs systématiques.

204. Il est essentiel de prévoir un temps suffisant pour un essai consciencieux des programmes d'ordinateur. Un moyen d'essai très efficace consiste à prendre un jeu de cartes perforées analogue à celui qu'on traitera ultérieurement, mais qui contient des erreurs et d'autres éléments délibérément insérés pour mettre en jeu toutes les particularités du programme. De cette façon, on peut l'essayer pour un prix modique et se servir des résultats pour le corriger si besoin est. En pratique, il est souvent impossible de préparer un jeu de cartes d'essai qui englobe toutes les possibilités.

205. L'informatique réduit beaucoup le volume du codage nécessaire et surtout du codage à main. En premier lieu, elle supprime les cotes dont l'objet principal est de faciliter le tri des cartes perforées. L'ordinateur peut traiter aussi bien une donnée numérique comme l'effectif des salariés que la cote à un seul chiffre qui lui correspond. De plus, si, pour d'autres raisons, on a besoin d'une telle cote, l'ordinateur peut aisément l'attribuer. Ensuite, il est notoire que les commis s'acquittent assez mal de l'attribution de cotes fondée sur le calcul ou sur la comparaison de grandeurs relatives, alors que l'ordinateur ne se trompe jamais, s'il a été programmé convenablement. En général, on préfère l'ordinateur chaque fois qu'une cote doit être attribuée à partir de données numériques, notamment si des calculs complexes interviennent. Les systèmes de codage peuvent aussi être simplifiés : la tabulation mécanique peut être facilitée par un numéro où la position de chaque chiffre a un sens (par exemple les deux premiers représentent la province; les deux seconds le district; et les deux suivants la ville); l'ordinateur opère tout aussi efficacement sur un numéro de classement à trois chiffres attribués aux villes, la référence au district et à la province étant conservée dans sa mémoire. Bien des pratiques adaptées au traitement manuel ou mécanique doivent être réexaminées si on veut recourir à l'informatique.

206. Toutefois, l'attribution par l'ordinateur d'une cote d'activité et de lieu exige une programmation extrêmement difficile et coûte très cher, car elle oblige à utiliser pendant un temps assez long un ordinateur d'une capacité suffisante pour mémoriser une masse de documentation de référence. Ces opérations ainsi que d'autres portant sur des renseignements non numériques sont le plus souvent mieux traitées par des commis.

Données provenant de dossiers administratifs

207. Les renseignements réunis sur les entreprises lors de la mise en oeuvre de divers programmes officiels offrent une possibilité d'utilisation tant comme source originale de données que comme instrument de collationnement. A cet effet, il est toutefois indispensable de disposer d'un moyen commode de transférer les données d'une source à l'autre. L'un d'eux consiste à attribuer des numéros matricules communs dans les dossiers administratifs et pour l'enquête. S'il est possible d'envisager de puiser les données dans les dossiers administratifs sans matériel mécanique ou électronique, l'exécution n'est guère pratique. Le matériel le plus indiqué est l'ordinateur qui peut exactement trier, assortir et transférer des données à grande vitesse.

208. Comme on l'a dit, pour puiser les données dans les dossiers administratifs au lieu de faire remplir un formulaire, il faut que l'administration utilise le même classement par genre d'activité et que ses dossiers comprennent les chiffres fondamentaux nécessaires à l'enquête, tels qu'emploi, salaires et volume des ventes, ainsi que toutes autres particularités nécessaires telles que forme juridique et lieu. A titre auxiliaire, mais encore important, ces dossiers administratifs peuvent servir de moyen de collationnement et de source de renseignements sur les enquêtés négligents. En tant que moyen de collationnement, ils fournissent une méthode autonome pour contrôler l'exactitude des données déclarées lors de l'enquête. Pour les établissements qui ont négligé de remplir le formulaire, ils servent de base pour interpoler les renseignements nécessaires quand les autres méthodes se révèlent inadéquates.

209. Etant donné que les techniques de traitement des données peuvent varier avec les services, il est souvent nécessaire d'adapter les bandes à diverses exigences. Cette adaptation peut demander beaucoup de temps et d'argent. Comme pour toute tâche d'informatique, il faut estimer la durée et le coût avant de décider.

Collationnement des enquêtes itératives par sondage

210. Etant donné que les enquêtes par sondage qui s'exécutent plusieurs fois par an obligent ordinairement à installer des services locaux répartis de façon appropriée dans tout le pays, ces services assureront mieux la collecte et le collationnement des déclarations que le service statistique central. En effet, ils pourront, pour résoudre les problèmes, prendre plus facilement contact avec les entreprises, par une visite, par téléphone ou par lettre. C'est là un avantage particulièrement important, car on tient d'ordinaire beaucoup à publier promptement les résultats de ces enquêtes, et il faut connaître vite les réponses aux questions.

211. Dans les cas où des enquêteurs sont chargés d'obtenir les renseignements lors de leurs entretiens, on peut les former à opérer simultanément de simples contrôles de collationnement et à inscrire toutes observations requises pour expliquer d'apparentes discordances. Elles seront examinées quand chacun remet son travail au service local. Quand le questionnaire doit être retourné par la poste, ce service peut exécuter une série de contrôles non seulement pour s'assurer de la concordance interne, mais encore pour comparer les données fournies lors d'enquêtes précédentes par la même entreprise. C'est là une particularité très commode des enquêtes itératives, car la même entreprise y répond régulièrement pendant une longue période, et on peut établir un canevas qui facilite la détection des erreurs.

212. L'objet principal de ce collationnement local dans les enquêtes qui ont lieu plusieurs fois par an est d'obtenir que les questionnaires soient remplis de façon correcte et exacte avant d'être transmis au service statistique central qui doit les dépouiller pour l'ensemble du pays. A ce stade, les formulaires ne subissent qu'un examen sommaire, car les tableaux sont préparés selon un calendrier très strict. Toute erreur découverte après mise en tableaux est difficile à corriger.

IX. MISE EN TABLEAUX, REVISION ET CORRECTION

Observations générales

213. Les types de tabulatrices normalement disponibles varient en complexité des simples machines de bureau à additionner aux grands ordinateurs rapides, fort onéreux, avec leur matériel auxiliaire spécialisé. Quand on a le choix du matériel, il est nécessaire de considérer toute une série de facteurs, comprenant le nombre de questionnaires à dépouiller, la quantité de données, le volume et le genre de collationnement à exécuter, le nombre et la complexité des opérations de tri et de tabulation, les délais prescrits et le coût des variantes. La décision doit être prise par des personnes qui connaissent parfaitement tous les matériels disponibles et qui comprennent bien les tâches à accomplir. Il convient d'étudier prudemment les avantages théoriques des ordinateurs et de consulter les pays qui leur ont confié des travaux analogues, avant toute décision. Les principaux matériels sont exposés ci-après.

Machines à additionner

214. Si les questionnaires sont en petit nombre et qu'on ne désire obtenir de sous-totaux que pour une ou deux caractéristiques des établissements, l'emploi des machines à additionner peut être plus rapide et plus efficace que toute autre méthode. A titre d'exemple, supposons que toutes les rubriques doivent être regroupées par a) genre d'activité et b) zone géographique. Après collationnement et codage, les questionnaires sont répartis de cette double façon. Les données pour chaque groupe de questionnaires correspondant à une branche d'activité dans une zone sont alors additionnées à la machine. Si les rubriques sont nombreuses, une machine enregistrant simultanément plusieurs colonnes facilite la tâche. On l'appelle quelquefois enregistreuse comptable. Elle doit être du modèle qui imprime, afin de permettre les vérifications. Après contrôle d'exactitude, les totaux peuvent être reportés sur des feuilles récapitulatives, qu'on additionne ensuite pour obtenir a) les totaux nationaux par genre d'activité et b) les totaux par région.

215. Si l'on a besoin d'autres tableaux - par exemple pour présenter tout ou partie des rubriques classées selon la taille de l'établissement -, il faut trier de nouveau les formulaires et recommencer les opérations. Le tri manuel exige d'ordinaire trop de temps si les tableaux de données doivent présenter plus de deux caractéristiques, et il faut alors envisager de tabuler par d'autres moyens. Il existe des dispositifs qui facilitent le tri, comme les cartes portant des perforations marginales, sur lesquelles on inscrit les données à la main ou à la machine à écrire. Ces dispositifs sont parfois utiles dans les enquêtes itératives mais, dans celles qui sont espacées, ils ne sont pas susceptibles de présenter plus d'avantages que les machines à additionner.

Traitement électromécanique

216. Le terme "électromécanique" sert à décrire le genre de matériel qui utilise des cartes perforées pour tabuler mécaniquement les données. Les

modèles principaux requis comprennent des perforatrices, vérificatrices, trieuses, collationneuses, reproductrices et machines comptables (tabulatrices). Avec ce matériel, les données provenant des questionnaires sont d'abord transcrites sur cartes perforées au moyen de la perforatrice, un autre opérateur vérifiant la perforation sur la vérificatrice. Les cartes sont alors triées et conférées en vue du tableau ou de la liste désirée; les tableaux sont dressés au moyen de la machine comptable.

217. Les matériels électromécaniques sont de vitesse et de capacité variables. La vitesse dépend en partie du nombre de sous-totaux et de totaux requis. Certains modèles sont capables de calculer et d'imprimer les totaux élémentaires, intermédiaires et principaux, pour le même ensemble de données : par exemple pour les provinces, les villes et tout le pays. Certains peuvent fournir de nouvelles cartes perforées donnant les totaux récapitulatifs pour des catégories déterminées. Ces cartes récapitulatives peuvent servir à des tableaux ultérieurs.

218. Après transcription des données, les cartes perforées peuvent soit passer dans des tabulatrices électromécaniques, soit alimenter des ordinateurs : directement s'il s'agit de petits ordinateurs, ou indirectement après transcription sur ruban magnétique. Passer les cartes perforées dans des trieuses et tabulatrices électromécaniques est particulièrement utile dans les recensements économiques où les données de base peuvent faire l'objet d'un classement simple ou double selon diverses caractéristiques, telles que genre d'activité, taille, lieu et mode d'exploitation de l'établissement. Si ces cartes alimentent des ordinateurs on obtient encore plus de souplesse, de rapidité et de précision.

219. Certaines machines ne peuvent perforer que des données numériques, tandis que d'autres traitent aussi les renseignements donnés en toutes lettres. Certains modèles impriment automatiquement chaque caractère en haut de la carte lors de sa perforation, ce qui peut aider à la vérification.

220. La perforation des cartes ne peut s'exécuter sans erreurs. Elles sont fréquemment plus importantes que celles qui se produisent dans la transcription manuscrite. Par exemple, s'il n'espace pas convenablement, l'opérateur peut facilement changer le nombre 123 pour en faire 123 000. De telles erreurs sont très nocives. Il est donc essentiel que toute la perforation fasse l'objet d'une vérification qui consiste à demander à un autre opérateur de perforer les données sur une vérificatrice. En général, toutefois, cette opération s'est révélée incomplètement efficace. On peut compter déceler environ 95 pour cent des erreurs si elle est d'efficacité moyenne.

221. Perforation et vérification seront normalement plus rapides et plus efficaces si les cartes sont conçues de façon que les chiffres y soient perforés dans l'ordre où ils figurent sur le questionnaire. Dans le traitement électromécanique, il est bon de réunir sur une même carte les rubriques à tabuler ensemble. Si, pour faciliter la perforation, des rubriques n'ont pas été perforées dans l'ordre correspondant à la tabulation,

il faut faire de nouvelles cartes au moyen d'une reproductrice, ce qui augmente le coût. S'il y a deux ou plusieurs cartes par questionnaire, le volume de renseignements aux fins d'identification et de classification à porter sur chaque carte dépend aussi de la méthode de tabulation. Avec le traitement électromécanique, il est d'usage de reporter sur chaque carte le numéro matricule de l'établissement, le codage par genre d'activité et par région et toutes les autres cotes, notamment celles concernant la taille, qui serviront à classer les données.

222. Avec les ordinateurs, on peut souvent, pour un coût faible ou nul, regrouper les ensembles de données quand on les fait passer des cartes perforées sur ruban magnétique. Si on se sert conjointement d'un enregistreur sur bande, il suffit d'enregistrer une fois la totalité des renseignements, ce qui n'est pas le cas avec le matériel électromécanique. L'ordinateur peut alors reconstituer électroniquement le questionnaire complet correspondant à l'établissement et lui donner toute autre disposition désirée.

223. Les progrès de la technologie ont permis de se passer entièrement des cartes perforées pour faire entrer les données dans l'ordinateur. Des deux principaux systèmes, l'un utilise le ruban magnétique (enregistreur sur bande) et l'autre accède directement à l'ordinateur (enregistreur sur disque). Un autre système se sert d'un "stylet électronique" au lieu d'une perforatrice pour transmettre directement les données à l'ordinateur. S'il n'est pas nécessaire de disposer d'abord d'un jeu de cartes perforées ou d'un ruban pour utiliser l'ordinateur, on peut, une fois qu'il a reçu les données, en préparer une bande, et il est également possible d'établir un jeu de cartes perforées à partir de cette bande.

Traitement informatique

224. Les progrès et usages des ordinateurs se sont multipliés au cours des vingt dernières années. Il en est résulté une technologie galopante où de nouvelles idées ont entraîné des changements continuels dans le matériel et le logiciel, avec un effet marqué sur le traitement des données. Toutefois, ces changements ont aussi provoqué une grande confusion quand sont apparues plusieurs générations successives d'ordinateurs et de matériels auxiliaires, ainsi qu'un certain nombre de langages correspondants. A plusieurs reprises, dans le présent manuel, on a averti des risques qu'il y aurait à se jeter dans l'informatique. Ce n'est pas pour nier sa vitesse, sa précision et sa souplesse supérieures, mais seulement pour prévenir que l'ordinateur ne peut que suivre ses instructions, dont la rédaction doit être complète et exacte dans tous ses détails. Selon l'étendue de sa mémoire et le genre de matériel auxiliaire, l'ordinateur peut traiter des masses de données, exécuter des classements et des calculs nombreux en exécution d'instructions complexes et imprimer les résultats dans la forme désirée. Pourvu qu'on dispose d'opérateurs, un préavis de quelques heures suffit pour commencer à tabuler, au moyen de machines à additionner, les données provenant d'un recensement industriel. L'emploi d'un matériel électromécanique demande plus de temps pour l'aménager et pour connecter les panneaux de commande. L'informatique en revanche, exige un préavis bien plus long. D'abord, il faut trouver

un programmeur entraîné qui devra prendre le temps de se familiariser avec les objectifs du recensement et en particulier avec les plans et prescriptions techniques concernant le collationnement et la tabulation. Cette condition peut obliger à élaborer des systèmes intensifs et extensifs avant de commencer à programmer. Ensuite, le programmeur doit esquisser son programme et en fixer les traits essentiels sous la forme d'un organigramme. Il en discutera alors en détail avec le personnel chargé de préparer l'enquête, en prenant tout le temps nécessaire pour s'assurer que chaque intéressé le comprend parfaitement. Il s'agit, par cette concertation, de découvrir tout malentendu et toute erreur d'interprétation qu'il serait extrêmement coûteux de corriger ensuite. On ne saurait trop insister sur la nécessité d'une initiation réciproque (qui, au programmeur, donne conscience des problèmes du statisticien et vice-versa) et d'une claire compréhension, problème capital de l'informatique.

225. Après approbation de son plan d'ensemble, le programmeur commence à rédiger le programme. Le temps nécessaire dépend de sa complexité et du genre d'ordinateur à utiliser. Les programmes simples peuvent ne demander que quelques jours; d'autres peuvent prendre des mois. Une fois écrit, le programme doit être essayé. On l'essaie généralement au moyen d'un jeu de cartes perforées analogue à celui qui servirait effectivement et conçu pour montrer les caractéristiques du programme et figurer les problèmes auxquels on s'attend dans le traitement effectif. Après essai et correction, le programme est prêt à servir pour collationner ou tabuler les données.

226. La vérification automatique de la perforation des informations codées est possible au moyen d'un numéro de contrôle, déterminé rapidement par l'ordinateur à partir de la cote proprement dite et ajouté à cette cote. A une cote à six chiffres, par exemple, le numéro de contrôle en ajoute un septième. Ce numéro peut se calculer de bien des façons. Dans un collationnement par ordinateur, il est recalculé et si le résultat obtenu ne concorde pas avec le résultat initial, la cote est signalée comme étant fausse.

Révision et correction des tableaux

227. L'efficacité du travail accompli pour concevoir le système de collecte des données et exploiter les questionnaires se vérifie véritablement quand les données sont rassemblées et présentées sous forme de tableaux. C'est seulement à ce stade qu'on peut juger de la plausibilité des résultats. Pour cela, il est nécessaire de disposer de moyens variés qui serviront, au personnel formé à leur emploi, à évaluer les tableaux.

228. Si un matériel perfectionné comme l'ordinateur permet des tests plus complexes, rien n'a remplacé encore le correcteur qui regarde un ensemble de chiffres et, en fonction de sa formation et de son expérience, juge s'ils sont plausibles. Toutefois, il ne doit jamais corriger arbitrairement au motif que certaines données sont fausses parce qu'elles paraissent invraisemblables. Il existe parfois des situations insolites, et les données doivent être vérifiées avec l'établissement qui les a fournies.

La révision des tableaux définitifs ne doit pas être une occasion de recommencer le travail exécuté lors des stades précédents. A ce stade-là, la révision consiste à examiner les données récapitulatives pour s'assurer de leur plausibilité. C'est seulement si le test en rejette une série suivie qu'on doit chercher à déterminer si la cause se trouve dans les opérations précédentes. Un budget approprié sera établi pour cette phase, car si on ne parvenait pas à discerner et à corriger les erreurs qui rendent les résultats inutilisables, tout l'effort consacré aux opérations précédentes serait perdu. La mauvaise exploitation d'un seul formulaire peut suffire à invalider les données pour une branche entière.

229. Deux genres principaux d'erreur peuvent causer des difficultés. Le premier consiste à inclure des données fausses (pour quelque cause que ce soit) pour un seul grand établissement ou pour un petit nombre d'entre eux : il se manifeste par un rejet casuel des données sérielles lors du test. Le second type d'erreur est systématique et entache de façon suivie une rubrique spécifique. Ce genre d'erreur peut apparaître même quand on ne révisé que les données concernant les petits établissements et risque de créer de grosses difficultés, car il peut obliger à recommencer une opération entière.

230. Les tableaux dressés à ce stade, quoique conçus pour comprendre tous les renseignements qui seront donnés dans la publication finale, n'ont pas besoin de revêtir la forme de cette publication. Au contraire, ils doivent être organisés essentiellement pour faciliter l'analyse. Leur conception et leur disposition matérielle doivent être conçues pour aider le réviseur à délimiter les données sérielles douteuses et pour guider autant que possible dans la recherche des causes.

231. Le collationnement des totaux peut s'opérer au moyen de deux sortes de vérification : par concordance interne et par confrontation avec des informations d'origine extérieure. Les contrôles de concordance interne cherchent à isoler les données sérielles qui sortent d'une fourchette acceptable lors de leur comparaison avec d'autres. Dans l'analyse des tableaux, ces fourchettes devront être bien plus étroites que pour un seul établissement, car on a maintenant affaire à une accumulation de données pour un groupe, où s'atténue beaucoup l'effet des cas extrêmes. En corollaire, plus le nombre d'observations sérielles est élevé, plus la fourchette doit être étroite et inversement. On dressera un barème des limites acceptables en fonction de ce nombre. A cet effet, on tiendra compte de facteurs tels que les variantes dues à la situation géographique et au genre d'activité. Si un ordinateur sert à préparer des tableaux pour révision, le barème sera fiché dans sa mémoire, de sorte que l'ordinateur reconnaîtra et signalera à l'attention les séries non conformes.

232. Le tableau qu'on trouvera le plus utile pour déceler les problèmes est celui qui présente les données dans le plus grand détail. Il s'agira d'ordinaire d'un tableau fournissant des données selon une subdivision détaillée des activités pour le plus petit secteur identifiable et donnant les renseignements de base sur le nombre des établissements, leurs ventes ou recettes, effectifs et salaires. Répercutées aux échelons supérieurs, la détection et la correction des erreurs importantes à cet échelon,

aboutiront à la présentation de tableaux acceptables à tous les échelons, pour autant que les mêmes rubriques y figurent. Dans la mesure où d'autres (par exemple nombre de chambres d'hôtel) ne figurent pas ailleurs, il faudra naturellement les revoir séparément.

233. Outre les soumettre aux contrôles de concordance interne, on peut comparer les données aux renseignements provenant d'enquêtes précédentes et d'autres sources. Une certaine difficulté, dans ce genre d'analyse, provient de l'évolution du niveau des prix et des taux de salaire avec le temps. Ils peuvent varier beaucoup d'un genre d'activité à l'autre et le réviseur devra avoir conscience de leurs différents taux d'évolution.

234. Dès lors qu'une série est reconnue douteuse, le réviseur doit trouver les moyens de déterminer l'erreur éventuelle. Si le cas semble systématique, il doit pouvoir remonter la chaîne des opérations jusqu'au point où a pu naître la difficulté. Pour cela, il doit être averti de tout ce qui s'est passé avant la mise en tableaux. Si le cas n'est pas systématique, il doit d'une certaine façon être capable de reconnaître les établissements qui ont pu le causer. C'est dire qu'il doit examiner la source initiale des données - le questionnaire - ou son équivalent. Pour être exploitable, le formulaire doit être disposé dans le même ordre que le tableau ainsi révisé, de préférence par secteur, par genre d'activité, par taille. Etant donné que le cas se borne presque toujours aux grands établissements, on peut borner à ce groupe le tri des formulaires.

235. Grâce aux ordinateurs, il est possible de dresser des listes de types divers qui peuvent contribuer à résoudre maints problèmes sans obliger à se reporter aux questionnaires. Ces machines peuvent rapidement répartir de gros fichiers en séries particulièrement utiles et éliminer la nécessité de consulter de volumineux dossiers où souvent les formulaires manquent ou sont mal placés. Ainsi, des listes peuvent être dressées et reliées en volumes pour permettre une consultation rapide par classes telles que secteur par genre d'activité, genre d'activité par taille mesurée selon divers critères, branche par secteur, etc. Même si on dispose de telles listes, il est parfois nécessaire de revoir les questionnaires.

236. Dès qu'une correction se révèle nécessaire, la question se pose de la meilleure façon de l'apporter. Quand une série de données ne figure qu'en un seul endroit, ce n'est guère difficile; mais, dans les tableaux de données économiques, il est courant que les données concernant un établissement figurent à différents échelons de présentation par secteur et genre d'activité et dans des classifications à double entrée. Il est donc indispensable de pouvoir apporter la correction nécessaire en plusieurs endroits des tableaux. Ces corrections peuvent être faites à la main ou bien portées sur les fichiers de cartes perforées ou de bandes à l'échelon soit de l'établissement soit de la plus petite carte récapitulative, tous les tableaux concernés étant alors refaits. Si l'on compte se servir des fiches d'établissement pour préparer d'autres tableaux dans un classement différent de celui des publications, il faudra corriger ces fiches, même si ce n'est pas nécessaire aux fins d'enquêtes pour lesquelles seules les cartes récapitulatives servent à préparer les tableaux définitifs.

237. On devra établir un document de référence unique où seront indiquées toutes les corrections à apporter en ce qui concerne l'établissement ou la série de données. A ce stade, les données nouvelles ou corrigées doivent être exactes, car elles serviront à préparer les tableaux définitifs qui seront publiés, et l'exécution des corrections doit se vérifier de façon absolument stricte. Aucune marge d'erreur n'est admissible.

238. Même si on se sert d'un matériel électromécanique ou d'ordinateurs, il est nécessaire de s'assurer que la somme des chiffres portés dans les colonnes est égale à leur somme portée en tête. Le nombre de discordances ainsi décelées étonne. La recherche de leur cause peut conduire à découvrir des erreurs graves.

239. Comme pour réviser les séries douteuses, le meilleur moment de rechercher les cas où la publication des données révélerait des renseignements sur l'exploitation d'un établissement ou d'une entreprise est celui où les données sont revues dans le plus grand détail. A ce stade, on se renseignera pour savoir si un seul établissement domine une série (cas, par exemple, de magasins d'alimentation dans une ville) à tel point que la publication des données permettra à des concurrents, à des fonctionnaires ou à d'autres d'avoir une idée très valable de ses recettes ou autres caractéristiques importantes. C'est à ce point qu'il devient particulièrement utile aux fins de comparaison de dresser la liste des principales entreprises selon l'ordre où elles apparaissent dans les tableaux.

240. La méthode ordinaire d'analyse de ce risque consiste à supprimer les séries dont les données, si elles étaient publiées, révéleraient les activités d'une unité statistique déterminée. Dans ce cas, l'unité est normalement l'établissement, mais la règle est souvent étendue aux activités des entreprises qui en possèdent ou contrôlent deux ou plusieurs. La publication de données pour une série ne comptant que deux unités statistiques est interdite, car elle révélerait à chacun l'activité de l'autre. Ainsi, on ne peut publier de données que pour les séries qui comprennent au moins trois unités statistiques.

241. Certaines statistiques sont considérées comme plus confidentielles que d'autres. Le montant des recettes et celui des salaires payés sont des exemples de données généralement tenues pour particulièrement critiques. Le nombre des établissements n'est pas d'ordinaire tenu pour confidentiel, à moins d'une classification si détaillée qu'elle révèle des renseignements confidentiels.

X. PUBLICATION

Révision finale

242. Avant leur dactylographie ou composition à l'imprimerie (ou avant phase finale des tableaux établis par ordinateur), les données doivent faire l'objet d'une révision finale. Ce travail exige un personnel tout particulièrement qualifié. Des commis et autres auxiliaires devront certes se charger des recherches nécessaires pour obtenir les renseignements requis par le réviseur, mais seul celui-ci devra déterminer les rubriques qui semblent douteuses et décider ce qu'il y a lieu de faire. Sa tâche n'est pas de déceler les erreurs minimales ni de contrôler de nouveau le collationnement, le codage et les corrections précédentes, mais d'examiner les données avec ses connaissances, son expérience et son intuition propres. S'il se révèle à ce dernier stade que les statistiques sont de mauvaise qualité, à cause d'erreurs soit dans les réponses soit dans l'exploitation ou que, s'il s'agit d'enquêtes par sondage, la variance qui en résulte est trop forte, mieux vaut supprimer les données que de les publier.

243. Dans cette opération finale, le réviseur doit avoir, autant que possible, des indications de l'ampleur de l'estimation et de l'interpolation qui ont contribué à obtenir les chiffres définitifs. En outre, il doit être averti de toutes les difficultés techniques qui ont été rencontrées et de la façon dont elles ont été résolues, de manière à pouvoir utiliser ces renseignements non seulement pour sa propre intelligence des statistiques, mais aussi pour préparer, en tant que de besoin, toutes notes explicatives ou mentions textuelles des réserves qu'appellent les données.

Analyse finale des risques d'indiscrétions

244. L'analyse directe des données pour découvrir les séries dont la publication révélerait ou permettrait aisément de découvrir des renseignements concernant l'exploitation de tel établissement aura été exécutée pendant la révision des tableaux. Il est un autre aspect de cette analyse qui doit s'opérer lors de la révision finale. C'est la recherche des révélations indirectes. Dans les tableaux statistiques à totaux horizontaux ou verticaux, une complication se produit si l'on supprime l'un des éléments sans supprimer au moins une autre donnée sérielle, car cet élément peut se retrouver en soustrayant du total la somme des éléments restants. Le risque de révélations directes sera d'autant plus grand que les tableaux renseigneront sur une plus grande quantité de détails; si on ajoute les révélations indirectes, il peut en résulter tant de suppressions que le tableau devient pratiquement inutile. Pour déterminer le degré de détail publié, on devra tenir compte du fait qu'à vouloir en donner davantage, on risque d'aboutir à diffuser moins de renseignements.

Préparation et reproduction des exemplaires

245. Les tableaux ayant été révisés et corrigés comme il convient, l'étape suivante consiste à leur donner leur forme définitive pour les faire paraître. Comme on l'a dit plus haut dans la section consacrée à la préparation des publications, le travail essentiel sur le contenu et la

forme des tableaux doit s'exécuter au début de la phase de préparation, car maintes opérations de l'enquête dépendent de ce qui sera publié.

246. La méthode de préparation et de reproduction des exemplaires est un facteur essentiel de détermination du temps nécessaire et du coût. Il en existe trois principales. La première consiste à compléter les tableaux à la main, par addition des rubriques, titres de colonne, titres généraux, notes explicatives, etc., après quoi la copie est envoyée à l'imprimeur pour composition, impression et assemblage. On obtient ainsi un document excellemment présenté, mais qui a l'inconvénient d'avoir coûté beaucoup de temps et d'argent. Dans la seconde méthode, les tableaux sont de même complétés manuellement, puis dactylographiés dans la forme exacte où ils seront publiés et reproduits par photo-offset. Le document obtenu, s'il présente moins bien, a l'avantage de coûter moins et de sortir plus vite. Dans la troisième méthode, le texte final est préparé par l'ordinateur. C'est la plus rapide et, en principe, la moins onéreuse, mais elle exige des programmes amples et difficiles et ne devrait être adoptée que si un tableau particulier compte assez de pages pour justifier le travail de programmation et le coût de l'informatique. Dans cette méthode, le texte se reproduit aussi par photo-offset, bien qu'on perfectionne actuellement les techniques pour composer par bande magnétique. Les textes sortant de l'ordinateur peuvent aussi être complétés manuellement et dactylographiés, pour ensuite être reproduits soit par impression, soit par offset.

247. Aucun mode de préparation des exemplaires n'est parfait; toutes les méthodes exigent un examen attentif pour éviter que des erreurs se produisent pendant la mise à jour manuscrite des tableaux, la dactylographie, la composition à l'imprimerie ou le passage par l'ordinateur.

Publication de rapports préliminaires

248. Le service statistique sera toujours fortement incité à publier les données dès que possible. Il peut de ce fait être tenté de diffuser des informations avant qu'on en ait convenablement vérifié la qualité. Si l'on constate ensuite des erreurs qui exigent d'importantes corrections, le service se trouvera soumis à d'abondantes critiques. Un compromis toujours utile, sinon complètement satisfaisant, consiste à sortir des rapports préliminaires qui donnent beaucoup moins de détails que n'en fourniront les documents définitifs et qui sont clairement présentés comme étant sujets à révision. Si c'est possible, les chiffres récapitulatifs destinés à cette communication préalable doivent représenter des totaux de contrôle qui ont été confirmés et qui se retrouveront dans les publications définitives.

Renseignements à publier au sujet des enquêtes

249. Pour permettre aux utilisateurs des statistiques publiées de déterminer leur comparabilité et d'en évaluer la qualité, les recommandations internationales adoptées par la Commission de statistique invitent à publier les renseignements suivants :

- a) Une description de la portée de l'enquête : préciser les types d'activité sur lesquels porte l'enquête et définir l'unité statistique eu égard au champ de l'enquête et aux catégories de données rassemblées.
- b) Une description du champ de l'enquête : préciser les limites éventuelles (en fonction de la dimension, du régime de propriété, du type d'organisation juridique, etc.).
- c) Une description des méthodes d'enquête : collecte directe des données (par correspondance ou sur le terrain), utilisation des dossiers administratifs ou sondages (préciser alors la technique de sondage utilisée et donner une estimation de la marge d'erreur imputable au sondage). Fournir éventuellement une estimation du degré de complétude du dénombrement.
- d) Les définitions opérationnelles des catégories de données rassemblées : fournir éventuellement des exemplaires des questionnaires et des instructions de base utilisés, y compris la description des méthodes utilisées pour déterminer la valeur des diverses catégories de données. Donner également la définition des catégories dérivées qui ont été construites à partir des catégories de données rassemblées.
- e) Le taux de non-réponse et les dispositions prises pour en tenir compte; préciser en particulier :
 - i) le nombre et l'importance des unités connues qui n'ont pas répondu au questionnaire et, si possible, certaines de leurs caractéristiques essentielles, c'est-à-dire leur type d'activité et surtout leur dimension. Indiquer également si les estimations relatives à ces unités ont été incluses dans les données publiées;
 - ii) le taux de non-réponse à des questions particulières pour lesquelles aucune estimation n'a été ou n'a pu être faite.
- f) Une description des classifications selon le type d'activité et des autres classifications utilisées.

250. En outre, il serait utile d'inclure dans le texte un exposé analytique des principales constatations de l'enquête. Cet exposé pourrait se fonder sur un jeu de brefs tableaux récapitulatifs présentant des rapports numériques (par exemple des stocks aux ventes, des salaires aux ventes, des salaires à l'effectif) par genre d'activité et région géographique, ainsi que des distributions en pourcentage (par exemple ventes réparties par établissement classé selon volume des ventes, effectif du personnel, région géographique). Ces tableaux pourraient être complétés

7/ Recommandations internationales pour les statistiques de la distribution et des services, par. 100.

par des graphiques qui rendent les statistiques plus frappantes et dégagent d'importants aspects de l'enquête.

Contenu des tableaux

251. Pour être le plus utiles, surtout à la présentation des données des recensements périodiques, les principaux tableaux devraient comprendre certains renseignements sur les périodes précédemment recensées. Ils figureraient dans les tableaux récapitulatifs donnant des statistiques par genre d'activité et par secteur. Il serait bon aussi que ces tableaux donnent les variations en pourcentage depuis le recensement précédent.

252. Pour les établissements se livrant à ces activités, l'utilité des renseignements concernant le commerce de détail et les services personnels est en raison directe de la petitesse du secteur pour lesquels ils sont publiés. Il en va particulièrement ainsi du commerce de détail. Les établissements ne s'intéressent pas tellement aux données qui se rapportent aux villes et aux provinces. Malgré la pression exercée pour obtenir des données relatives aux principaux quartiers marchands et centres commerciaux, cette tâche ne doit pas être entreprise sans une idée claire de l'effort, du temps et de l'argent qu'elle exigera.

253. Un problème particulier aux tableaux concernant les services provient du caractère hétérogène de ces derniers. Il est encore aggravé par le fait que la recommandation sur la portée à leur donner tout en déclarant "qu'on a établi la liste des services énumérés en choisissant les activités les plus susceptibles d'être pratiquées dans tous les pays", ajoute cette réserve que "ces services n'ont pas tous la même importance du point de vue du rassemblement des données, encore qu'on puisse difficilement leur attribuer à chacun une priorité distincte. Il peut se révéler nécessaire, pour un pays donné, d'ajuster la portée des enquêtes en supprimant par exemple des services peu importants ou non existants et en incorporant ceux qui présentent une certaine importance dans ledit pays". Il ressort clairement de ces remarques que, aux fins de comparabilité internationale, il faut examiner soigneusement la portée donnée aux totaux correspondant aux services selon les pays. De même, les genres d'activité réunis sous cette rubrique - agences immobilières, producteurs de films cinématographiques, hôtels, agences de publicité, émissions de radio, blanchisseries, informatique, etc. - en rendent le total difficile à saisir. Fondamentalement, les données pour de nombreux services n'ont de sens que si on prend chaque catégorie séparément. Cette situation oblige fréquemment à réunir des données spécialisées pour la catégorie et à préparer les tableaux correspondants. Quand on a réuni suffisamment de ces renseignements particuliers, la publication d'un rapport consacré à la catégorie peut se justifier.

254. Les tableaux contenant des données qui proviennent en tout ou en partie de sondages doivent clairement l'indiquer. Ces données doivent toujours figurer après multiplication par le coefficient approprié. Le fait que les totaux obtenus à partir de sondages ne concordent pas avec ceux provenant d'un dénombrement intégral doit être expliqué et on doit toujours donner des estimations de la variance due au sondage.

255. Presque tous les tableaux à publier présenteront des données de base selon un classement à double entrée d'après les caractéristiques des établissements. Le nombre de tableaux à préparer de cette façon est si grand que chaque pays devra en choisir un petit nombre offrant le plus grand intérêt général. Des exemples en sont donnés par les titres et notices ci-dessous. Les tableaux considérés se fondent sur la liste de statistiques à tabuler selon le genre d'activité figurant dans les recommandations internationales. La liste convient mieux pour tabuler les résultats de recensements, car il n'est pas recommandé de réunir certaines de ces rubriques lors des enquêtes annuelles ou qui ont lieu plusieurs fois par an. La liste n'est pas limitative et les titres n'y sont pas forcément rangés par ordre d'importance. La présentation des données par région géographique, au-delà de ce que propose le tableau 1, sera déterminée par chaque pays en fonction de ses besoins et des ressources dont il dispose.

Tableau 1. Principaux indicateurs d'activité, classés par groupe de la CITI et par région

Ce tableau présenterait les rubriques essentielles, telles que nombre d'établissements, effectif occupé, nombre de salariés, salaires et traitements payés au personnel, montant des ventes de marchandises et des recettes d'exploitation, valeur ajoutée et valeur de la formation brute de capital fixe. Les groupes de la CITI pourraient être ceux énumérés au paragraphe 62 des recommandations internationales. La classification territoriale correspondrait aux grandes divisions administratives du pays, telles qu'Etats, provinces, etc. Si on a besoin de détailler plus largement, il faudra sans doute réduire la classification par activité aux classes (trois chiffres) et catégories (deux chiffres) de la CITI.

Tableau 1a. Commerce de gros et de détail : principaux indicateurs d'activité, classés par groupe de la CITI et par région

Ce tableau présenterait, pour le seul commerce de gros et de détail, les mêmes rubriques fondamentales que le tableau 1 plus la marge brute, classées par groupe de la CITI et par région. Les groupes de la CITI pourraient être ceux proposés au paragraphe 62 des recommandations internationales ou une classification nationale comparable. Comme pour le tableau 1, chaque pays devra considérer le degré de détail de la classification géographique en le rapportant à celui de la classification par activité, de façon à disposer d'un tableau utilisable et de longueur maniable.

Tableau 2. Principaux indicateurs d'activité, classés par catégorie de la CITI et par taille d'établissement

Ce tableau présenterait les mêmes rubriques fondamentales que le tableau 1, dans un classement à double entrée par genre d'activité selon les catégories (deux chiffres) de la CITI et par taille d'établissement. Pour définir la taille, les critères évidents sont le nombre de personnes occupées et la valeur totale des ventes et des recettes.

Tableau 2a. Commerce de gros et de détail : principaux indicateurs d'activité, classés par groupe de la CITI et par taille d'établissement

Ce tableau présenterait, pour le seul commerce de gros et de détail, les mêmes rubriques fondamentales que le tableau 1a, dans un classement à double entrée par genre d'activité selon les groupes de la CITI (quatre chiffres) proposés au paragraphe 62 des recommandations internationales ou selon une classification nationale comparable et par taille d'établissement.

Tableau 3. Principaux indicateurs d'activité, classés par catégorie de la CITI et par type d'établissement

Ce tableau présenterait les mêmes rubriques fondamentales que le tableau 1, dans un classement à double entrée par genre d'activité selon les catégories (deux chiffres) de la CITI et certains genres d'organisation économique et juridique et de propriété. Si ces genres sont nombreux, il serait préférable en pratique de diviser ce tableau en deux ou trois parties, à raison d'une par caractéristique.

Tableau 3a. Commerce de gros et de détail : principaux indicateurs d'activité, classés par groupe de la CITI et par type d'établissement

Ce tableau présenterait, pour le seul commerce de gros et de détail, les mêmes rubriques fondamentales que le tableau 1a; elles figureraient dans un classement à double entrée par genre d'activité selon les groupes (à quatre chiffres) de la CITI proposés au paragraphe 62 des recommandations internationales ou une classification nationale comparable et par type d'organisation économique et juridique selon des genres choisis, mode d'exploitation et type de propriété. Si ces genres sont nombreux, il serait préférable en pratique de diviser le tableau en deux, trois ou quatre parties, à raison d'une par caractéristique.

Tableau 4. Indicateurs de l'emploi et des gains, classés par groupe de la CITI

Ce tableau présenterait pour l'ensemble du pays des rubriques sur l'emploi et les gains, tels que nombre de propriétaires exploitants et travailleurs familiaux non rémunérés, effectif des salariés (en distinguant si possible les permanents), salaires et traitements payés aux salariés (en distinguant si possible ceux payés aux permanents) et les compléments aux salaires et traitements. Les données seraient classées selon les groupes (à quatre chiffres) de la CITI ou l'équivalent national.

Tableau 5a. Commerce de gros et de détail : valeur des ventes et recettes, classée par groupe de la CITI

Ce tableau présenterait, pour le seul commerce de gros et de détail, des données sur le produit brut total et ses éléments constitutifs, telles que valeur des ventes de marchandises réparties entre ventes en gros (à subdiviser encore en ventes pour compte propre, ventes pour le compte d'autrui et commissions reçues) et ventes au détail, et total des recettes d'exploitation. Les données seraient classées selon les groupes (quatre chiffres) de la CITI proposés au paragraphe 62 des recommandations internationales.

Tableau 5b. Commerce de gros et de détail : valeur des ventes, classée par genre de marchandise

Ce tableau présenterait, pour le seul commerce de gros et de détail, des données sur la valeur des ventes de marchandises réparties entre gros et détail par genre de marchandise ou groupes de marchandises. Chaque pays devra dresser des listes spéciales pour la déclaration des ventes par genre de marchandise. Il serait bon que celle des marchandises vendues au détail soit liée à la classification des biens et services ménagers servant à la comptabilité nationale 8/.

Tableau 6. Valeur des achats et autres dépenses d'exploitation, classée par groupe de la CITI

Ce tableau présenterait des données sur la valeur du capital circulant et de ses éléments constitutifs, telle que montant des achats de marchandises (celles destinées à la vente et les autres), électricité et combustible, travaux de réparation et d'entretien, traitances et commissions et autres coûts d'exploitation. Les données seraient classées selon les groupes (quatre chiffres) de la CITI ou l'équivalent national.

Tableau 6a. Commerce de gros et de détail : valeur des achats de marchandises destinées à la vente, classée par genre

Ce tableau présenterait, pour le seul commerce de gros et de détail, des données sur la valeur des achats de marchandises destinées à la vente dans ces mêmes commerces, par genre de marchandise ou groupes de marchandises.

8/ Les études préparées par le Bureau de statistique des Nations Unies, au cours de l'établissement d'une classification générale des marchandises, fourniront un guide pour élaborer ces classifications. Voir "Classification internationale type de tous les biens et services" (E/CN.3/493).

Tableau 7a. Commerce de gros et de détail : valeur des stocks, classée par groupe de la CITI

Ce tableau présenterait, pour le seul commerce de gros et de détail, la valeur des stocks en début et fin d'exercice, en distinguant si possible ceux de marchandises destinées à la vente. Figureraient aussi dans ce tableau le mouvement des stocks en cours d'exercice et leur renouvellement. La classification se ferait selon les groupes (quatre chiffres) de la CITI proposés au paragraphe 62 des recommandations internationales.

Tableau 8. Formation brute de capital fixe et sa composition, classé par groupe de la CITI

Ce tableau présenterait des données sur la valeur des acquisitions d'actifs fixes et additions brutes à ces actifs, par genre (c'est-à-dire bâtiments et autres édifices, terrains, matériel de transport, machines et autres matériels). La classification se ferait selon les catégories (quatre chiffres) de la CITI ou l'équivalent national.

Enquêtes ayant lieu plusieurs fois par an

256. Pour publier les données provenant de telles enquêtes, la ponctualité présente bien plus d'importance que pour les enquêtes plus espacées, mais le volume très limité des données rend moins ardues la préparation et la reproduction des exemplaires. Les données provenant d'un sondage minimisent le risque d'indiscrétions mais tendent aussi à rendre la révision finale plus difficile, car il est plus malaisé de retrouver la cause des erreurs de dépouillement. Le texte doit comprendre une description du sondage et un tableau de ses variances à l'intention des lecteurs. Etant donné qu'ordinairement, le temps presse pour la publication, il est normal que le texte présente une analyse sommaire des données sous une forme telle qu'elle puisse être reproduite dans sa totalité ou utilisée avec quelques retouches.

Annexe I

TECHNIQUES DE SONDAGE DANS LES ENQUETES SUR LA DISTRIBUTION ET LES SERVICES

1. Il est souvent souhaitable de sonder quand on a besoin de statistiques sans nécessiter un recensement complet. D'ordinaire en effet, un sondage coûte bien moins qu'un tel recensement et dérange bien moins les enquêtés. Certes, il engendre l'erreur de sondage, mais la tâche étant plus légère peut s'exécuter avec plus de soin, ce qui réduit éventuellement l'erreur totale. Le sondage permet aussi une publication bien plus prompte des résultats.
2. Dans le sondage probabiliste, chaque unité comprise dans l'univers statistique a une probabilité non nulle d'être choisie et le choix final est opéré par le hasard et non par le jugement de quiconque. S'il est convenablement conduit, l'erreur de sondage peut se calculer à partir de l'échantillon. Les échantillons sont souvent tirés d'autres façons, par exemple en prenant les unités qu'on peut obtenir le plus aisément ou celles qui, au dire d'un tel, sont représentatives de l'univers statistique. L'erreur de sondage ne peut pas être calculée dans ces cas, ce qui réduit beaucoup leur intérêt. Le reste de l'exposé se borne donc au sondage probabiliste.
3. Etant donné que les échantillons, dans ces sondages, sont tirés au hasard, de nombreux résultats s'écarteront plus ou moins de la moyenne (valeur escomptée). Cet écart à la moyenne s'appelle erreur de sondage. La variance se définit comme la moyenne des carrés de ces écarts. La racine carrée de la variance s'appelle erreur type. Dans l'hypothèse ordinairement valable d'une forme normale, on peut connaître la distribution des erreurs à partir de l'erreur type. Le quotient de l'erreur type par la grandeur considérée s'appelle coefficient de variation. Cette mesure de l'erreur relative permet de comparer l'erreur de sondage entre résultats très différents par la taille et le caractère. Le carré de cette mesure s'appelle la variance relative, qui est égale au quotient de la variance par le carré de la grandeur. Le calcul de l'erreur de sondage décrit ci-dessous se fait à partir de la variance, qu'on peut aisément ramener aux autres paramètres mentionnés, si on le désire.
4. On réduira l'erreur de sondage en augmentant la taille de l'échantillon (en gros, l'erreur type varie en raison inverse de la racine carrée de cette taille). Une démarche souvent suivie dans la considération de l'erreur de sondage pour les échantillons de tailles différentes consiste à calculer (ou à supposer empiriquement) une variance relative pour une seule unité de l'échantillon, puis à diviser cette valeur par le nombre proposé d'unités dans l'échantillon pour obtenir la variance relative correspondante. Par exemple, si on suppose empiriquement égale à 1 la variance relative par établissement commercial compris dans un sondage sur liste, la variance relative d'un échantillon comprenant 100 établissements sera 0,01, et le coefficient de variation sera de 0,1 ou 10 pour cent. (Noter que ni le nombre d'unités de l'univers statistique ni leur taille moyenne n'interviennent dans ce calcul).

5. Ces enquêtes sont exposées à un autre genre d'erreur, dite erreur non imputable au sondage, qui ne résulte pas du choix au hasard et qui, en général, ne dépend pas de la taille de l'échantillon. Cette erreur, par exemple, peut provenir de défauts de la base utilisée, d'erreurs de déclaration, de l'absence de déclarations pour certaines unités, d'erreurs de dépouillement, etc. L'erreur de ce type est particulièrement dangereuse, car elle ne se prête pas à la mesure comme l'erreur de sondage. Il faut la contenir dans des limites acceptables par des programmes appropriés de contrôle qualitatif, même si on doit sacrifier la taille de l'échantillon pour y parvenir. Pour décider du compromis à établir entre les efforts à consacrer à réduire respectivement l'erreur de sondage et l'erreur non imputable au sondage, on doit se souvenir qu'il s'agit de parvenir à la moindre erreur totale possible avec les ressources disponibles, et non pas de minimiser la seule erreur de sondage.

6. Il découle de ce principe qu'un plan de sondage simple qu'on peut exécuter comme prévu est préférable à un plan plus élaboré, supérieur en théorie, mais que sa complexité même expose à de graves erreurs non imputables au sondage dans son exécution.

7. Le budget de toute enquête (et la taille de l'échantillon qui lui correspond) doit théoriquement se fixer au point où le coût de l'enquête, augmenté des séquences de décisions fautives résultant d'erreurs statistiques, passe par un minimum. Il est d'ordinaire impossible de déterminer ainsi ce budget, même grossièrement, surtout à cause de la difficulté d'évaluer les séquences de décisions fautives. En outre, souvent, les budgets ne dépendent pas des statisticiens. Toutefois, la considération du principe énoncé ci-dessus peut empêcher de commettre des erreurs grossières, comme par exemple d'opérer un recensement complet quand une enquête par sondage suffirait ou de prendre un échantillon trop petit.

8. Pour opérer un sondage, la première démarche consiste à décider de la base d'où sera tiré l'échantillon. Dans le cas d'établissements commerciaux, cette base est une liste, soit de tous ces établissements, soit des unités auxquelles on peut les associer (comme la liste des entreprises propriétaires ou la liste des lieux où ils sont situés).

9. Les bases servant aux sondages des établissements commerciaux peuvent se répartir en deux grandes catégories :

- a) Liste des établissements ou des entreprises qui les possèdent (souvent tirée des rôles fiscaux ou des registres des associations professionnelles). Les sondages opérés sur ce genre de base sont appelés ci-après sondages ou échantillonnages sur liste.
- b) Une liste des localités du pays. Pour éviter la distorsion, toutes doivent avoir une chance d'être choisies. Les sondages opérés sur ce genre de base sont appelés ci-après sondages ou échantillonnages sur superficie (ou encore aréolaires).

On peut aussi opérer des sondages ou échantillonnages mixtes si une partie de l'univers statistique est sondée sur liste et une autre partie sur superficie.

10. Pour l'erreur qu'il entraîne, le sondage sur liste est d'ordinaire préférable à l'autre, car :

- a) le sondage par superficie oblige à une tournée pour mettre à jour l'inscription des établissements commerciaux dans les localités choisies. Dans le sondage sur liste, l'opération peut souvent se faire essentiellement par lettre ou par téléphone, ce qui revient ordinairement moins cher;
- b) le sondage sur liste se stratifie d'ordinaire plus efficacement;
- c) le sondage sur superficie est un sondage par grappes qui réduit d'ordinaire l'efficacité par cas choisi. Le sondage sur liste n'est pas ordinairement par grappes.

11. Si le sondage sur liste l'emporte nettement d'ordinaire du point de vue de l'erreur de sondage, il pose de graves problèmes pour l'erreur non imputable au sondage. Il est souvent très difficile de donner à chaque établissement commercial une possibilité et une seule d'être choisi sur une liste, car :

- a) la liste n'est jamais à jour en ce qui concerne les établissements tout nouvellement créés;
- b) elle omet d'ordinaire, par définition, une partie de l'univers statistique considéré. Habituellement, mais non toujours, ce sont les petits établissements qui sont omis;
- c) la liste omet aussi quelques établissements qui, par définition, devraient y figurer. Par exemple, les fraudeurs ne figurent pas sur les rôles fiscaux; ceux qui n'en sont pas membres ne sont pas inscrits sur les registres des associations professionnelles;
- d) la liste contient souvent des doublets : un même établissement peut faire l'objet de deux ou plusieurs inscriptions qui ne sont pas exactement identiques par le nom, l'adresse ou tout autre moyen d'identification;
- e) dans un sondage sur liste, des établissements qui n'ont pas fourni de déclaration peuvent être portés d'office comme étant en activité, alors qu'en fait, ils sont fermés, ou inversement;
- f) de même, il est courant de rayer de la liste d'adresses les unités de sondage déclarées comme ayant cessé leur activité. Cette déclaration peut être fausse ou périmée; les unités qui avaient cessé leur activité au moment de la déclaration ont pu la reprendre. Ce facteur tend à provoquer une distorsion cumulative, cause de sous-estimation.

12. En général, le sondage sur superficie, parce qu'il se fonde sur des observations concrètes, fournit une base solide pour juger exactement

si un établissement commercial est en activité ou non et pour donner à chacun une chance et une seule d'être choisi. Toutefois, son champ se trouve aussi exposé au risque de distorsion. Les principales difficultés sont les suivantes :

- a) Certains établissements commerciaux ne se manifestent pas par des signes extérieurs. En général, l'enquêteur les saute.
- b) L'enquêteur se trompe aussi dans l'établissement de sa liste. En général, il inscrit moins d'établissements qu'il ne devrait. Un contrôle qualitatif peut d'ordinaire minimiser ces erreurs.

13. On notera que les erreurs d'inclusion seront plus souvent cumulatives dans les sondages sur liste que dans les autres. De telles erreurs sont particulièrement gênantes, car elles faussent les estimations aussi bien des variations que des quantités.

14. Le choix d'une base pour une enquête donnée doit se faire en fonction de la source qui entraînera la moindre erreur totale avec le budget disponible; ce qui dépend des conditions particulières de l'enquête considérée. Toutefois, les principes directeurs suivants pourront se révéler utiles :

- a) Dans une large mesure, le choix dépend de la qualité des listes disponibles. Si des omissions ou doublets les déparent, mieux vaudra probablement choisir un sondage aréolaire. A noter (voir d) ci-dessous) qu'il n'est pas recommandé de combiner un sondage sur liste incomplète et un sondage aréolaire pour remédier aux insuffisances de la liste.
- b) On évitera les bases formées par des listes multiples qui se recoupent. Théoriquement, on peut tenir compte de la probabilité de doublets qui en résulte, si elle est connue, pour les cas sondés, mais il est pratiquement très difficile de déterminer dans quelle mesure elle existe.
- c) Si on dispose d'une seule liste satisfaisante ou d'un groupe de listes qui ne se recoupent pas et dont la réunion forme une liste satisfaisante, le choix est moins clair :
 - i) pour une enquête qui n'a lieu qu'une seule fois (par exemple un sondage à l'occasion d'un recensement), le sondage sur liste s'impose;
 - ii) ce même genre de sondage peut aussi convenir aux enquêtes périodiques espacées si on dispose de beaucoup de temps (par exemple six mois ou plus) pour tabuler. Il en va ainsi parce qu'on dispose du temps nécessaire pour résoudre les problèmes de champ qui se posent;
 - iii) les plus épineux de ces problèmes se rencontrent avec les enquêtes fréquentes (par exemple mensuelles ou trimestrielles) qui ne laissent que peu de temps (par exemple un mois ou

moins) entre leur fin et la préparation des tableaux. Quasi certainement, le sondage sur liste posera ici un grave problème de champ. Sera-t-il plus important que l'erreur de sondage plus accusée qu'entraîne l'échantillonnage aréolaire ? La réponse dépend en partie du montant du budget disponible. Etant donné que l'erreur non imputable au sondage opéré sur liste ne diminue pas avec la taille de l'échantillon, elle peut représenter une plus forte partie de l'erreur totale dans une enquête à gros budget;

iv) en partie, la décision dépend de la possibilité de collecter les déclarations par la poste ou par téléphone. Sinon, il pourra être beaucoup plus coûteux de visiter les unités d'un sondage sur liste assez dispersées que celles d'un sondage aréolaire par grappes.

d) Si on recourt au sondage aréolaire, il est recommandé de le compléter par une liste des grands établissements ou sociétés intégralement dénombrés. Etant donné que ces unités d'inclusion certaine font l'objet d'une tournée chaque mois, il est facile de suivre leur position de façon à les "sauter" comme il convient dans le sondage. En outre, la probabilité d'une double inscription exigerait une double déclaration de l'enquêté. Il n'est pas recommandé de combiner sondages sur liste et sondages aréolaires quand la liste n'est pas entièrement dénombrée. Le problème consiste ici à choisir les unités à sauter parmi celles de l'échantillon aréolaire qui ont une chance d'être choisies. Il est d'ordinaire si difficile que cette combinaison n'est pas recommandée.

15. Les paragraphes suivants proposent une méthode pour tirer un échantillon aréolaire d'établissements commerciaux. La première démarche consiste à décider du caractère de la grappe finale, définie comme un groupe de m segments d'échantillon aréolaire contenant chacun une moyenne de \bar{N} établissements commerciaux et inclus dans une unité primaire de sondage (UPS). Au sujet de la grappe finale, il y a lieu de décider :

- a) la taille de \bar{N} ;
- b) la taille de m ;
- c) la taille et la nature de l'unité primaire de sondage.

16. Les réponses à ces questions dépendent du coût et de la variance de l'enquête considérée. Toutefois, les recommandations empiriques suivantes devraient aboutir à des résultats suffisamment proches de l'optimum dans la plupart des cas :

- a) \bar{N} , nombre des établissements commerciaux à retenir dans chaque segment, doit être le plus proche possible de 6 (4 si on ne retient que les magasins de détail). Cet optimum a été trouvé par investigation empirique dans les conditions existant aux

Etats-Unis d'Amérique et peut servir de point de départ dans d'autres pays. Toutefois, étant donné que le coût et la variance peuvent différer beaucoup dans ces pays, il est recommandé d'entreprendre des recherches empiriques pour déterminer la taille qui assure la plus grande fiabilité par unité de dépense.

- b) Le nombre de ces segments retenus dans chaque UPS destinée à être sondée doit être aussi proche que possible de m , c'est-à-dire du nombre de segments dont peut aisément s'occuper un seul enquêteur.
- c) La taille et la nature des unités primaires de sondage devraient se déterminer par les considérations suivantes :
 - i) pour chacune, on devrait disposer d'un étalon de taille, par exemple recettes des établissements commerciaux, ou estimations démographiques provenant d'un recensement précédent;
 - ii) l'UPS ne doit pas être si grande qu'elle oblige l'enquêteur à passer le plus clair de son temps à aller de son domicile au segment ou d'un segment à l'autre. (Le trajet moyen de son domicile au segment ne devrait peut-être pas dépasser trois quarts d'heure, si c'est possible);
 - iii) l'UPS ne devrait pas être trop petite, sous peine d'une grande dégradation due à une forte corrélation intraclasses;
 - iv) dans ces limites, l'UPS devrait être aussi hétérogène que possible (par exemple, chacune devrait combiner des zones urbaines et rurales si possible).

17. Les UPS seront probablement des subdivisions administratives ou des combinaisons de celles-ci. Si un sondage aréolaire des ménages est opéré dans le pays, très probablement le même jeu d'UPS peut servir au sondage aréolaire des établissements commerciaux. Il n'est toutefois pas recommandé que l'échantillon des segments ayant servi à cette enquête serve au sondage des établissements commerciaux, car la corrélation entre ménages et établissements commerciaux est très faible à ce niveau.

18. La démarche suivante consiste à déterminer l , nombre des grappes finales qui figureront dans l'échantillon. Ce nombre se détermine en divisant le budget disponible à cet effet par le coût estimatif par grappe finale. Le montant de ce budget doit se déterminer d'après les considérations exposées au paragraphe 7 ci-dessus. La démarche suivante consiste à envisager l'erreur de sondage probable (coefficient de variation ou CV) pour les échantillons aréolaires des diverses tailles proposées. On calcule d'abord \sqrt{ml} , c'est-à-dire la racine carrée du nombre de segments dans le plan proposé. On divise l par ce nombre pour obtenir une valeur approximative du CV des quantités; 0,2 par ce nombre pour obtenir une valeur approximative du CV du mouvement annuel. Par exemple, soit une enquête

comptant 40 grandes grappes finales (1) de 10 segments (m) chacune en moyenne. On compte ainsi 400 segments (ml). Le CV est alors de 5 pour cent

($\frac{1}{20}$) pour les quantités, de 1 pour cent ($\frac{0,2}{20}$) pour le mouvement mensuel et de 2 pour cent ($\frac{0,4}{20}$) pour le mouvement annuel. Ces approximations pourront

être affinées quand on disposera de données provenant d'essais préalables ou d'enquêtes effectives. A partir de l, on peut calculer le taux global de sondage P, par la formule $P = 6lm/N$, où N est le total estimatif des établissements commerciaux que compte le pays. P étant ainsi fixé, et les échantillons étant tirés à ce taux, ne doit plus changer, même si N se révèle erroné.

19. Déterminer toutes les UPS d'inclusion certaine dans le pays. Ce sont celles pour lesquelles $S_j \geq \frac{S}{I}$, formule où S_j est un étalon de taille, S est

le même étalon à l'échelon national et l est le nombre de grappes finales comprises dans l'échantillon, défini ci-dessus. Par exemple, supposons que le nombre d'habitants serve d'étalon de taille pour tirer les UPS. Si ce nombre (S) est estimé à 10 millions et qu'on doive tirer 100 grappes finales (l), toute UPS qui compte plus de 100 000 habitants ($\frac{S}{I}$) doit être incluse.

20. Stratifier toutes les UPS d'inclusion incertaine restantes. Les strates doivent être d'une taille aussi proche que possible de $2Sl$, c'est-à-dire qu'elles sont fonction de l'étalon S utilisé. Dans notre exemple, on s'efforcerait d'obtenir des strates comptant 200 000 habitants. Les UPS placées dans une strate donnée devraient être aussi analogues que possible.

21. Diviser en deux sous-strates chaque strate formée comme il est dit au paragraphe précédent. Ces sous-strates doivent être d'une taille aussi égale que possible (ici, la taille de référence sera S/l , soit 100 000 habitants).

22. Dans chaque sous-strate ainsi formée, choisir une UPS avec une probabilité proportionnelle à S. Poursuivant notre exemple, supposons une sous-strate comprenant trois UPS : A avec 10 000 habitants, B avec 30 000 et C avec 60 000. Nous calculons comme suit un étalon cumulatif de taille :

<u>UPS</u>	<u>Etalon de taille</u>	<u>Etalon cumulatif de taille</u>
A	10 000	10 000
B	30 000	40 000
C	60 000	100 000

Choisissons ensuite un nombre aléatoire entre 1 et 100 000. Pour tout nombre entre 1 et 10 000, on choisit l'UPS A; pour tout nombre entre 10 001 et 40 000, l'UPS B; pour tout nombre entre 40 001 et 100 000, l'UPS C. Les probabilités de choix sont ainsi respectivement de 10 000/100 000 (ou 0,1), de 30 000/100 (ou 0,3) et de 60 000/100 000 (ou 0,6).

Note : Cette méthode de choix donne une UPS sondée par strate. En conséquence, il est impossible d'obtenir une estimation sans distorsion de la variance. On en obtient une estimation globale par la méthode dite de la demi-strate, exposée ci-dessous. Elle suppose que les deux UPS proviennent de la même strate.

23. On dispose de certains schémas pour choisir deux UPS par grande strate avec une probabilité proportionnelle à la dimension et sans remplacement. Toutefois, ces schémas sont plus complexes et leur variance effective est plus forte que celle calculée dans le système décrit ici.

24. Une autre variante possible consiste à tirer deux UPS par grande strate avec une probabilité proportionnelle à la dimension et avec remplacement. Cette méthode fournit de la variance une estimation sans distorsion, mais la variance effective est supérieure à celle calculée par la méthode décrite ci-dessus.

25. Après détermination des UPS d'inclusion certaine et choix des autres UPS, la démarche suivante dans le sondage aréolaire consiste à subdiviser en segments toute la superficie des UPS choisies. Chaque segment doit présenter les caractéristiques suivantes :

- a) des limites suffisamment permanentes (rues, routes, voies ferrées, cours d'eau, etc.) reconnaissables par les enquêteurs;
- b) un nombre d'établissements commerciaux aussi proche que possible de 6. Si le plus petit segment possible aux limites reconnaissables en compte 10 ou davantage, il doit être dit "d'inclusion partielle", la mesure de sa dimension s'obtenant en divisant par 6 le nombre estimatif des établissements (quotient arrondi au chiffre le plus proche).

26. Pour opérer cette segmentation, il est nécessaire de disposer de cartes où figurent des limites reconnaissables telles que cours d'eau, routes, etc., et l'emplacement des établissements commerciaux. Si ces cartes existent déjà (à la suite par exemple d'un recensement précédent des établissements commerciaux), on s'en servira. Il n'est pas nécessaire qu'elles soient absolument précises ou à jour. Si elles n'existent pas, il faut les établir par une tournée personnelle des UPS sondées. Partant de cartes aussi détaillées que possible (agrandies si besoin est), les enquêteurs devront dresser la liste de tous les établissements commerciaux (déterminés par observation extérieure) et ajouter toutes caractéristiques reconnaissables qui ne figurent pas déjà sur les cartes.

27. Grâce aux cartes déjà existantes ou établies comme indiqué au paragraphe précédent, l'UPS sondée sera segmentée au bureau statistique conformément aux règles du paragraphe 25. Cette segmentation doit englober toute la superficie de l'UPS sondée.

28. Les segments de l'UPS sondée (par inclusion certaine ou par choix) seront numérotés à la suite (par exemple en zig-zag). Dans cette numérotation, chaque mesure d'un segment "d'inclusion partielle" sera traitée comme un segment séparé.

29. On calcule alors les taux interne de sondage. Pour les UPS d'inclusion certaine, il est égal à P , c'est-à-dire le taux global de sondage. Pour les autres, il est égal à $Pgh_j = P \cdot Sgh/Sgh_j$.

30. Dans cette dernière formule, S est une mesure de dimension utilisée dans le sondage des UPS, Sgh mesure la sous-strate h comprise dans la i ème (grande) strate et Sgh_j mesure la j ème UPS comprise dans cette strate. La fraction Sgh/Sgh_j est donc l'inverse de la probabilité de choix de l'UPS considérée.

31. Pour poursuivre notre exemple, supposons qu'on ait choisi l'UPS A : la fraction Sgh/Sgh_j est alors égale à $1/0,1$, soit 10. Avec un taux global de sondage (P) de $1/100$, le taux interne pour l'UPS A sera décuple, soit $1/10$. L'échantillon sera effectivement choisi en prenant pour intervalle l'inverse de la probabilité interne de choix, soit $1/P$ ou $1/Pgh_j$, avec un point de départ aléatoire situé entre 0 et ce nombre. Pour choisir les segments, le canevas qui en résulte est appliqué aux segments numérotés à la suite dont il est question au paragraphe 28. Les segments de chaque UPS d'inclusion certaine sont systématiquement assignés, sitôt tirés, à cinq groupes aléatoires pour faciliter le calcul des variances exposé au paragraphe 63.

32. Il est bon d'ordinaire de compléter le sondage aréolaire par une liste des très grands établissements (ou entreprises). Cette liste peut s'obtenir d'une source quelconque, mais tous les établissements qui y figurent doivent être inclus dans le sondage et dénombrés pour chaque période d'enquête. La limite d'inscription sur cette liste sera initialement fixée à $S/6ml$, soit S divisé par le total estimatif des établissements compris dans l'échantillon aréolaire. S est une mesure de dimension (par les ventes, l'effectif, les salaires) disponible aussi bien à l'échelon national que pour chaque établissement important. Il n'est pas nécessaire que ce soit la même qui sert à sonder l'UPS. Par exemple, si le total des ventes au détail dans un pays est estimé à un milliard de dollars et qu'on escompte 2000 cas sondés, la limite sera fixée à $1\ 000\ 000\ 000/2\ 000$, soit 500 000 dollars; et on s'efforcera d'inscrire sur la liste d'inclusion certaine tous les établissements ayant au moins cette taille.

33. C'est là une approximation très grossière, qui suppose essentiellement que l'augmentation de la variance et celle du coût unitaire pour les établissements soumis au sondage aréolaire (à la différence de ceux du sondage sur liste) s'annulent dans la formule optimale de répartition. Une limite plus exacte se détermine de façon empirique, par tâtonnement, quand on a établi quelques listes en vue du sondage aréolaire. Dans ces essais, on doit égaliser le coût total en faisant varier la taille de l'échantillon aréolaire, dans la mesure nécessaire pour compenser la variation de taille de l'échantillonnage sur liste.

34. L'enquêteur doit être invité à inscrire sur sa liste tous les établissements commerciaux qui se trouvent à l'intérieur des limites des segments, avec les exceptions suivantes :

- a) Sauter tous les établissements inscrits sur la liste d'inclusion certaine. Pour ce faire, l'enquêteur reçoit une liste de ces établissements (ou entreprises) dont on sait qu'ils opèrent, ou pourraient opérer, dans son UPS.
- b) Pour les établissements "d'inclusion partielle", ceux qui figurent sur la liste doivent être sondés selon la fraction k/K , où K mesure la taille attribuée au segment et k est le nombre (d'ordinaire égal à 1) des "cas" sondés qui se trouvent dans le segment. L'échantillonnage peut être opéré sur place par l'enquêteur, ou au bureau local. S'il est fait sur place, le point de départ et l'intervalle doivent être désignés d'avance et les règles données à l'enquêteur pour l'ordre d'inscription ne lui laisseront aucune latitude. L'échantillonnage au bureau local, moins exposé au risque de distorsion, peut exiger un important surcroît de temps et d'argent.

35. Un sous-échantillon du listage et du dénombrement exécutés par chaque enquêteur sera soumis à un contrôle de qualité. Ce contrôle est plus efficace s'il est opéré séparément, c'est-à-dire si l'enquêteur et le contrôleur préparent des listes séparées et les collationnent. Il est aussi plus efficace s'il n'est pas fait selon un calendrier fixe qui permettrait à l'enquêteur de savoir quels mois seront contrôlés.

Note. Les questions de pondération, de calcul de la variance, d'utilisation des renseignements auxiliaires dans les méthodes d'estimation, et de rotation s'appliquent aux sondages aussi bien sur liste que sur superficie et sont exposées aux paragraphes 58 à 95.

36. Si on décide de sonder sur liste, la première opération consiste à stratifier cette liste, c'est-à-dire à répartir les établissements qui y sont inscrits en classes (strates) d'où seront tirés des échantillons indépendants. Il s'agit de rendre les unités de sondage comprises dans les strates aussi semblables que possible; les critères convenables sont la catégorie d'activité et la taille, si on les connaît. Pour chaque catégorie, cinq ou six strates de taille, d'inclusion incertaine, devraient procurer la plupart des avantages, sans être trop complexes. Toutefois, si la stratification incombe à l'enquêteur au cours de sa tournée, il est recommandé pour simplifier de s'en tenir à une seule strate d'inclusion incertaine. L'objectif est d'égaliser autant que possible la taille de ces strates. Toutefois, pour simplifier, on peut décider d'utiliser la même limite de taille pour chaque catégorie d'activité. Les strates seront ainsi de taille égale à l'échelon de l'ensemble des catégories d'activité, mais inégale dans chacune de ces catégories.

37. La taille totale de l'échantillon sur liste se déterminera en divisant le budget total disponible (coûts fixes exclus) par le coût de l'unité de sondage. Le montant total du budget devrait se déterminer d'après les considérations exposées au paragraphe 7. En envisageant l'erreur de sondage, on peut supposer à l'origine que le coefficient de variation d'une quantité estimée à partir d'un sondage sur liste sera égal à 1 divisé par

la racine carrée du nombre d'unités de sondage. Le coefficient de variation du mouvement mensuel pourrait être proche du quotient de 0,3 par cette racine carrée; celui du mouvement annuel, du quotient de 0,5 par cette même racine carrée (pour utiliser l'exemple du paragraphe 18 sur la façon de faire ces calculs, prendre le nombre d'unités à tirer à tirer dans le sondage sur liste). Si la catégorie d'inclusion certaine est importante, opérer ces calculs pour la catégorie d'inclusion incertaine et multiplier les résultats par PNC, c'est-à-dire la proportion des recettes totales due à cette dernière catégorie. Ces grossières approximations initiales peuvent être affinées quand les données provenant de l'enquête deviennent disponibles.

38. Une méthode simple pour répartir l'échantillon global entre les diverses catégories d'activité consiste à proportionner cette répartition à la taille globale de chaque catégorie. On obtient ainsi un total général quasi optimal. L'erreur de sondage (coefficient de variation) pour chaque estimation catégorielle sera à peu près inversement proportionnelle à la racine carrée de la taille. Certes, si on pense que certaines catégories d'activité sont plus ou moins importantes que ne l'indiquerait cette répartition, on peut la modifier.

39. L'attribution optimale à une catégorie d'activité dans les diverses strates de taille est donnée par la formule :

$$n_h = \frac{N_h \sigma_h}{\sum_{n=1}^R N_n \sigma_n} n$$

où n_0 désigne l'échantillon total attribué à cette catégorie d'activité (conformément au paragraphe 38 ci-dessus);

n_h désigne l'échantillon attribué à la h ème des R strates de taille dans cette catégorie;

σ_h désigne l'écart type (pour un seul tirage) de la h ème strate.

Si l'attribution indiquée par cette formule est égale ou supérieure au nombre N_h d'unités de la strate, toutes ces unités seront incluses.

40. Une approximation simple et souvent satisfaisante de l'attribution ci-dessus s'obtient ainsi :

- a) Fixer la limite d'inclusion certaine à X/n_0 , où X est la mesure de la taille totale pour la catégorie d'activité et n_0 l'échantillon attribué à cette catégorie.
- b) Attribuer le reste de l'échantillon proportionnellement à X_h (mesure de la taille totale dans la h ème groupe de taille) dans la catégorie considérée. Par exemple supposons qu'un

échantillon de 100 établissements a été attribué à une catégorie dont la taille globale (par exemple montant des salaires payés) est de 100 millions de dollars. Tout établissement versant au moins un millions en salaires sera inclus. S'il y en a dix versant au total 20 millions, il reste à répartir un échantillon de 90 établissements d'inclusion incertaine. A raison de cinq strates, chacune en recevra 18.

41. L'approximation ci-dessus suppose des coefficients de variation (pour un tirage unique) égaux pour toutes les strates. En fait, ils sont d'ordinaire plus forts pour les classes de moindre taille. Une attribution un peu meilleure s'obtient parfois en majorant légèrement chaque mesure de taille avant d'opérer comme il est dit aux paragraphes 39-40 ci-dessus. Par exemple, si la taille se mesure à l'effectif occupé, on peut le majorer d'une unité pour chaque établissement (ou entreprise, si elle sert d'unité de sondage) avant de répartir l'échantillon.
42. Pour tirer l'échantillon indiqué dans chaque strate de taille par catégorie d'activité, inscrire d'abord et numéroter systématiquement toutes les unités de sondage de la strate (par taille, catégorie d'activité, emplacement, etc.). Puis choisir un échantillon au moyen du canevas obtenu en utilisant un intervalle $I_n = \frac{N_h}{n_h}$ (arrondi à l'entier le plus proche) et un point aléatoire compris entre 1 et I_h . Les unités de sondage seront attribuées systématiquement, sitôt tirées, aux cinq groupes ou redoublements aléatoires, en vue des calculs de variance décrits plus loin.
43. Comme on l'a dit, les principales difficultés du sondage sur liste en vue d'une enquête périodique consistent à faire en sorte que chaque établissement ait une chance et une seule d'être choisi et à décider sans se tromper quels établissements compris dans le sondage sont effectivement en activité. Dans les paragraphes suivants sont présentées des recommandations en vue de les atténuer.
44. On tiendra un registre de toutes les unités qui ont une chance d'être choisies à un moment ou à un autre. Avec ce registre, on vérifiera si les unités qui viennent de se créer et celles qu'on croit nouvelles ont déjà eu une telle possibilité.
45. On devrait disposer d'une bonne indication des créations d'établissements : ces établissements seraient sondés et ajoutés à l'échantillon dès que possible.
- a) Si, au moment de la création, on ne dispose pas d'informations valables en vue de la stratification (par catégorie d'activité et éventuellement par taille), on sondera à deux degrés. A cet effet, on tire d'abord un gros échantillon pour obtenir ces informations. De cet échantillon on en tire un autre plus petit, de façon que le produit de la première probabilité par la seconde convienne pour la strate finale. La taille de l'échantillon du premier degré (c'est-à-dire la fraction de sondage des créations d'établissements) ne se prête pas à une optimisation

mathématique explicite. Comme dans les autres problèmes d'optimisation de ce type, on peut procéder par essais empiriques dès que deviennent disponibles les données de l'enquête. Ces essais doivent s'opérer comme suit :

- i) à partir des cas sondés, recréer un univers statistique en pondérant chaque cas par l'inverse de la probabilité de choix. Cet univers comprendra à la fois les unités de l'échantillon original et les créations d'établissements;
- ii) essayer pour ces créations diverses fractions de sondage au premier degré dans cet univers simulé, Noter que les fractions les plus fortes donneront moins d'aberrances (causées par les cas dont la probabilité au premier degré est moindre que celle qui convient à leur strate finale). En revanche, elles majoreront le coût, à cause de l'agrandissement de l'échantillon;
- iii) pour chacune de ces fractions de sondage au premier degré, calculer le coût correspondant. Ajuster la taille de l'échantillon global en fonction des différences de ces coûts, de façon que le coût total reste constant;
- iv) calculer la variance globale correspondant à chaque fraction en se servant des probabilités indiquées et choisir le plan qui assure l'équilibre optimal entre les effets des aberrances précitées et la taille de l'échantillon.

A moins qu'on ne dispose des données provenant de l'enquête, les probabilités initiales de ce sondage au premier degré doivent être assez fortes pour qu'elles ne risquent pas d'entraîner d'aberrances. Par exemple, s'il se peut que des créations soient d'une taille qui justifie l'inclusion certaine, le sondage au premier degré doit s'opérer de cette façon. Un tel sondage rendra aussi plus efficaces les essais précités, puisque l'univers ainsi simulé se rapprochera de l'univers réel.

- b) Deux systèmes principaux servent au sondage des créations. Dans l'un, on ne sonde que les créations authentiques (c'est-à-dire les établissements sans prédécesseurs); les successeurs sont représentés par les successeurs des unités déjà comprises dans le sondage. Dans l'autre système, on sonde toutes les créations et on élimine tous les établissements qui cessent leur activité, qu'ils aient un successeur ou non. Etant donné qu'on n'a pas de définition nette et claire des successeurs et qu'il n'est pas toujours facile d'obtenir des renseignements sûrs à leur sujet, ce second système est probablement moins sujet à distorsion cumulative.

46. Il est recommandé de s'enquérir de la situation des enquêtés qui n'ont pas fourni de déclaration pour savoir s'ils sont toujours en activité. De

cette façon, on décidera sans se tromper s'il faut leur attribuer une estimation non nulle. Il est aussi recommandé de s'enquérir, au moins une fois par an, des fermetures, pour voir si les établissements n'ont pas repris leur activité.

47. Il est recommandé de recommencer fréquemment le tirage pour les raisons suivantes :

- a) réduire l'effet des distorsions cumulatives;
- b) répartir la charge des déclarations;
- c) éviter les accroissements de variance dus à des changements dans les mesures de taille.

Un nouvel échantillon peut être indépendant du précédent, ou le recouvrir plus ou moins. Tirer un échantillon indépendant est de loin le plus simple et assure la plus grande rotation possible.

48. Tirer un nouvel échantillon sur liste provoquera probablement une discontinuité entre anciens et nouveaux résultats. C'est là l'un des inconvénients du sondage sur liste dans les enquêtes périodiques. Certes, un échantillon aréolaire devra aussi faire l'objet d'un nouveau tirage, mais moins souvent et, comme il est moins sujet aux distorsions cumulatives, il risque moins d'entraîner de discontinuités marquées.

49. Tirer un échantillon d'un recensement complet représente une autre application des principes du sondage sur liste aux établissements commerciaux. Ce tirage doit s'opérer (pour exiger moins de temps et d'argent tant des enquêtés que de l'Etat) quand certaines données n'ont pas besoin d'être trop détaillées territorialement (par exemple recettes par catégorie de marchandises, stocks, dépenses d'équipement). Cette application évite bien des problèmes exposés aux paragraphes 43 à 48 ci-dessus, car un recensement complet comprenant des rubriques telles que le nombre d'établissements et les recettes totales donne, de l'ensemble statistique considéré alors, une liste qui est relativement à jour, sans omissions ni doublets. Toutefois, elle pose quelques problèmes particuliers, qui dépendent de la méthode de démarchage utilisée pour le recensement et qui sont exposés ci-après.

50. Il existe trois méthodes générales d'opérer un recensement complet des établissements commerciaux. La première consiste à organiser des tournées d'enquêteurs chargés de dresser par dénombrement la liste des établissements à recenser. La seconde consiste à enquêter (d'ordinaire par correspondance ou téléphone) à partir d'une liste préétablie d'établissements commerciaux souvent tirée de dossiers administratifs. Une troisième méthode consiste à tirer directement des dossiers administratifs les données de base souhaitées. Le recensement peut s'opérer en combinant ces trois méthodes, chacune s'appliquant à des parties différentes du champ. Les problèmes de sondage posés par chacune des trois sont exposés ci-dessous.

51. Si un recensement s'opère en tout ou partie par tournée d'enquêteurs sans liste préalable, l'échantillon d'établissements sera tiré pendant ou après la tournée. Théoriquement, on peut le tirer avant par sondage aréolaire, mais il ne semble guère y avoir de raison d'accepter les défauts d'un tel sondage quand on dispose, grâce à la tournée exécutée, d'une liste à jour. Si l'échantillon est tiré en cours de dénombrement, il doit être extrêmement simple et ne laisser à l'enquêteur aucune latitude de jugement personnel. D'ordinaire, il est bon d'avoir deux strates, l'une d'inclusion certaine et l'autre d'inclusion incertaine. Par exemple, la règle serait de sonder tous les établissements occupant au moins dix personnes et 1 sur 10 de ceux qui en occupent moins (cette limite et la fraction de sondage seraient déterminées par les principes exposés aux paragraphes 36 à 42). A l'enquêteur seraient prescrites des règles inflexibles pour déterminer les établissements à sonder dans la seconde strate. Par exemple, si l'échantillon est préétabli sur des listes, on lui prescrirait un itinéraire à suivre exactement. Variante, chaque questionnaire pourrait porter sous cachet la désignation de l'échantillon. Ce cachet ne serait rompu que le recensement terminé. L'enquêteur apprendrait alors si l'établissement appartient ou non à l'échantillon.

52. En dépit de ces précautions, un échantillon tiré par un enquêteur en cours de dénombrement subira presque certainement une distorsion, encore que souvent faible et n'invalidant pas les résultats. L'effet de cette distorsion (ou celle causée par d'autres facteurs) peut souvent se réduire par la procédure d'estimation exposée au paragraphe 57 ci-dessous.

53. Au lieu de faire tirer l'échantillon par l'enquêteur au cours de sa tournée, on peut le tirer après la fin du recensement de base. Dans ce cas, l'enquêteur envoie les questionnaires aux services nationaux ou régionaux, où ils sont sondés (probablement à réception). Cette méthode permet d'utiliser des systèmes de stratification plus complexes et plus efficaces. Elle réduit aussi le risque de distorsion. Toutefois, elle peut se révéler inutilisable pour deux raisons :

- a) Elle accroît le temps total nécessaire pour opérer le recensement, car elle introduit un délai entre le recensement de base et le sondage. Il s'ensuit non seulement un retard dans la publication, mais encore un risque de distorsion, car un nombre important d'établissements peuvent cesser leur activité dans l'intervalle.
- b) Elle accroît le coût total, car il faut opérer un dénombrement (par tournée d'enquêteurs, correspondance ou téléphone) pour compléter le sondage. Si l'enquêteur sonde pendant le dénombrement de base, il peut ramasser tout ou partie des questionnaires servant au sondage ou du moins les laisser aux enquêtés.

54. Si tout ou partie d'un recensement se fait par dénombrement (probablement par lettre ou téléphone) sur une liste préétablie, on peut tirer l'échantillon soit avant d'expédier le questionnaire de base, soit

après l'avoir reçu en retour. Dans chaque cas, le sondage s'opère dans le service, de sorte qu'on peut appliquer des plans plus complexes et plus efficaces fondés sur les principes énoncés aux paragraphes 36 à 42. Toutefois, si le sondage s'opère avant l'envoi des questionnaires de base, on risque de manquer de bons renseignements, nécessaires pour stratifier, sur la taille et le genre d'activité. Il en résulte un problème particulièrement difficile si le questionnaire à expédier pour le sondage dépend du genre d'activité.

55. Si le sondage s'opère après le retour des questionnaires de base, on dispose pour stratifier de meilleurs renseignements sur la taille et le genre d'activité. Ici encore, la méthode entraîne un important surcroît de temps et de coût, car le dénombrement en vue du sondage doit suivre (peut-être avec un intervalle important), celui en vue du recensement de base.

56. Le recensement de base opéré en tout ou partie par compilation de dossiers administratifs offre encore le choix de tirer l'échantillon avant ou après réception des données. Souvent, ce n'est qu'après qu'on dispose d'une liste à jour à cet effet. Souvent aussi, il est alors trop tard pour pouvoir échantillonner ainsi. Il n'y a donc pas lieu à réunion de données pour cette partie du champ, surtout si elle est faible.

57. Quelle que soit la façon de sonder sur une liste, on disposera des données de base (par exemple nombre d'établissements et recettes totales) tant pour la partie recensée que pour la partie sondée. Ces données peuvent servir à obtenir de meilleures estimations que celles provenant du sondage, par application d'un coefficient fractionnaire de la forme $\frac{Y'}{X'} X$. Dans cette expression, Y' désigne une estimation d'une certaine rubrique provenant du questionnaire de sondage; Y représente une grandeur statistique en corrélation provenant du recensement de base; X' représente la même grandeur estimée d'après le sondage. (La grandeur X doit être choisie en corrélation particulièrement étroite avec l'estimation Y'). Par exemple, supposons que Y' représente l'estimation des ventes d'une certaine catégorie de marchandises (comme le rayon vêtements dans les grands magasins) et X représente le total des ventes dans ces magasins. On obtient d'ordinaire une meilleure estimation (Y'') des ventes de vêtements par la formule ci-dessus. Cette formule offre quatre avantages :

- a) L'estimation fractionnaire Y'' peut avoir une moindre variance que l'estimation Y' si la corrélation est suffisante entre X et Y' .
- b) L'estimation peut servir de moyen d'interpoler les réponses qui manquent pour Y' (sondage), dans l'hypothèse qu'il n'en manque pas pour X (recensement).
- c) Le procédé peut servir à atténuer l'effet des distorsions provoquées lors du sondage du fait de l'enquêteur ou d'autres causes. Il suppose que la fraction $\frac{Y'}{X}$ est la même pour la

population statistique exacte que pour le sondage imparfait. Ce n'est pas nécessairement vrai, mais souvent plus près de la vérité que de supposer l'estimation Y' exempte de distorsion.

- d) Si les recettes totales sont subdivisées (par exemple par catégorie de marchandises), la somme de ces éléments redonnera les recettes totales indiquées par le recensement de base (pourvu qu'on pose $X' = Y'$ en totalisant tous ces éléments).

Ces trois derniers avantages ne justifient pas qu'on emploie le procédé pour cacher les défauts notables d'un sondage. Si la différence entre les valeurs X' (du sondage) et les valeurs X (du recensement) dépasse ce qu'on peut normalement imputer à l'erreur de sondage, les utilisateurs doivent être informés des fractions $\frac{X}{X'}$, utilisées, de façon à connaître le degré

de correction des distorsions et, éventuellement, d'interpolation, opérée par ce procédé.

58. Les paragraphes suivants exposent quelques principes des sondages, tant sur liste que sur superficie. Le premier point concerne la pondération. Les autres sont successivement le calcul des variances, la rotation des échantillons, l'emploi de renseignements auxiliaires pour réduire les variances et les dangers d'une réduction arbitraire de pondération.

59. On pondère toujours par l'inverse des probabilités de choix pour obtenir des estimations simples exemptes de distorsion. Dans le cas de sondages aréolaires, cette probabilité est le produit des probabilités aux divers degrés. S'il s'agit d'un sondage à deux degrés, elle s'obtient en multipliant la probabilité de tirer l'unité primaire de sondage par la probabilité de tirer les segments contenus dans cette unité.

60. Dans l'exemple donné aux paragraphes 29 à 31, l'UPS a été choisie avec une probabilité de $1/10$ et les segments de cette UPS avec une même probabilité du $1/10$. Ce n'est pas par coïncidence que la probabilité totale de $1/100$ est égale à la fraction de sondage préétablie (P) du $1/100$. Le coefficient de pondération à utiliser est alors de 100. Si le tirage s'opère comme il est recommandé, la probabilité de choix (P) est uniforme pour toutes les parties du sondage, et les résultats sont pondérés par $1/P$. (Un sondage où les coefficients de pondération sont uniformes est souvent appelé auto-pondéré).

61. Dans le cas d'un sondage sur liste, le coefficient de pondération est l'intervalle (I_h) servant au choix (voir paragraphe 42), ou l'unité s'il y a inclusion certaine.

62. Les estimations simples ainsi pondérées et exemptes de distorsion servent toujours soit d'estimations finales, soit de termes pour obtenir les estimations plus complexes exposées ci-dessous. Elles servent aussi au calcul de la variance, soit directement, soit comme termes si l'on utilise l'un des procédés plus complexes d'estimation.

63. Pour les plans de sondage (sur liste comme sur superficie) recommandés ici, la variance de l'estimation simple sans distorsion peut se calculer par la formule suivante, qui convient à tout sondage stratifié :

$$\sum_{n=1}^R (1-f_h) \frac{n_h}{m_h-1} \sum_{j=1}^{n_h} (Y'_{hj} - \bar{Y}'_h)^2 \quad (1)$$

Dans cette formule, h désigne la hième strate (sur un total de R). Dans le sondage aréolaire, chaque UPS d'inclusion certaine forme une strate séparée. Pour les autres UPS, chacune des strates originales d'où sont tirées deux grappes finales forme une strate, que ces grappes soient tirées en créant d'abord deux sous-strates, ou directement de la strate originale (avec ou sans remplacement). Dans le sondage sur liste, chaque catégorie de taille-genre d'activité d'où est tiré un échantillon forme une strate; f_h est la fraction de sondage, c'est-à-dire la probabilité (ou inverse du coefficient de pondération) utilisée dans la hième strate. Avec la méthode de la demi-strate (c'est-à-dire avec un seul tirage par sous-strate), f_h est nul. Y'_{hj} désigne la valeur simple pondérée du j^{ème} redoublement dans la hième strate. Dans le sondage à un seul degré, ce paramètre désigne la valeur pondérée du j^{ème} groupe de tirages. Dans le sondage à plusieurs degrés, il désigne la valeur pondérée de la hième grappe finale dans la hième strate, que le tirage s'opère sur la strate ou (comme dans la méthode de la demi-strate) à raison d'une grappe finale sur chacune des deux sous-strates.

Le nombre de groupes mentionnés ci-dessus est de m_h dans la hième strate. Dans la partie d'inclusion incertaine du sondage aréolaire, il est égal à 2 si le tirage s'opère comme recommandé \bar{Y}'_h désigne la moyenne simple des paramètres Y'_{hj} susmentionnés soit :

$$\bar{Y}'_h = \frac{\sum_{j=1}^{m_h} Y'_{hj}}{m_h}$$

Si on se sert d'une estimation plus complexe que l'estimation simple sans distorsion, la variance se calcule selon cette même formule où U'_{hj} remplace Y'_{hj} . Cette expression U'_{hj} désigne les combinaisons linéaires de Y'_{hj} . Elles sont données pour les estimations particulièrement complexes indiquées ci-dessus. Pour connaître une méthode générale de calcul, le lecteur se reportera à la bibliographie : cette méthode a fait l'objet d'un programme d'ordinateur qu'on peut obtenir du Bureau de recensement des Etats-Unis d'Amérique. Ce programme calcule les variances d'estimations complexes à partir des seules données requises pour obtenir l'estimation elle-même.

64. Les sondages servant aux enquêtes périodiques sur les établissements commerciaux font souvent l'objet d'une rotation dans le temps. Deux raisons la justifient :

- a) répartir la charge des réponses entre des enquêtés plus nombreux;
- b) réduire l'erreur de sondage entachant les estimations.

65. Le renouvellement périodique des sondages sur liste, recommandé pour réduire la distorsion cumulative (paragraphe 47 et 48) peut de même, sans obliger à rotation, mieux répartir la tâche entre les enquêtés. Pour avoir cet effet, en revanche, les sondages aréolaires risquent d'être trop peu fréquents, car l'erreur cumulative y pose un moindre problème. Il faudra alors recourir à la rotation.

66. La rotation du sondage opérée pour réduire la variance est assez complexe et expose donc au risque d'une distorsion supplémentaire. Il est recommandé d'adopter un cadre fixe si on a le moindre doute sur l'aptitude à maîtriser ces complexités. On trouvera ci-après les principes essentiels d'un plan de rotation. Pour plus de détails, se reporter à la bibliographie.

67. Il est recommandé d'utiliser à chaque fois un nombre limité de mêmes cadres rotatifs (appelé un ensemble fermé). On s'assure ainsi deux avantages sur l'emploi d'un nouveau cadre pour chaque nouvelle période :

- a) Préparer un cadre de dénombrement coûte cher, que ce soit sur liste ou sur superficie. Limiter le nombre de cadres limite le coût initial.
- b) Avec un ensemble fermé, on peut appliquer la méthode de grande observation, qui peut contribuer notablement à réduire la variance.

68. Pour une enquête mensuelle, il est commode d'utiliser un ensemble fermé de 12 cadres. C'est un nombre commode parce qu'il lui permet d'avoir un cadre pour janvier, à dénombrer en janvier de chaque année, un pour février, et ainsi de suite jusqu'en décembre. De cette façon, les estimations pour chaque mois de l'année se fonderont sur le même ensemble de cadres d'une année à l'autre.

69. Les importantes estimations du mouvement annuel s'en trouveront améliorées, à cause de la corrélation qui existe d'une année à l'autre. Si le coût initial est trop élevé avec 12 cadres, on peut en prendre moins. Avec 2, 3, 4 ou 6, les mêmes reviendront chaque année; avec 3, le même jeu reviendra chaque trimestre. Certes, réduire le nombre de cadres réduit quelque peu les avantages qu'ils offrent, mais même un ensemble réduit à 4 assure à la variance l'essentiel de l'amélioration qu'apporte la rotation.

70. Deux calendriers différents de déclarations peuvent servir avec un ensemble de cadres rotatifs. Si de chaque enquêté à chaque tournée on peut obtenir deux mois de données, tout le cadre doit être remplacé lors de la période suivante. Par exemple, prenons le cadre de janvier dans l'ensemble de 12 auquel il appartient. Début février, lors de la tournée si on peut obtenir de l'enquêté qu'il renseigne à la fois sur janvier et

on ne conserve que celui de février pour la tournée suivante. Si, en revanche, on ne peut pas obtenir début février de renseignements fiables sur les ventes de décembre, on ne demandera que les ventes de janvier avec le cadre du même mois. Toutefois, pour obtenir les statistiques de ventes provenant d'établissements identiques, on conservera ce cadre pendant un mois de plus et on s'enquerra de leurs ventes de février lors de la tournée du cadre de février début mars. Avec ce plan, chaque cadre tourne à raison de deux fois par an au lieu d'une fois. Si l'on veut que l'effort demandé soit le même que selon le premier plan, chaque cadre peut être de dimension moitié moindre.

71. Manifestement, le premier plan fournit plus de renseignements par unité de sondage. En l'appliquant, on peut fonder la mesure du mouvement mensuel pour des établissements identiques sur l'ensemble du cadre et non sur la moitié. Toutefois, ce plan expose à un risque de distorsion additionnelle qui n'est pas négligeable. Les déclarations obtenues pendant le mois en cours doivent indiquer les mêmes valeurs que les déclarations faites le mois précédent. Prenons de nouveau janvier pour exemple. On obtiendra deux déclarations au sujet de l'activité pendant ce mois. L'une (l'immédiate) sera obtenue du cadre de janvier dès la fin de ce mois. Une autre estimation pour le mois de janvier s'obtiendra grâce au cadre de février, au début de mars. Plusieurs raisons peuvent entraîner des différences entre les valeurs ainsi déclarées successivement. Un, si les déclarations sont tirées des registres, ceux de janvier seront probablement plus complets au début de mars qu'au début de février, ce qui fera déclarer des montants différents pour les ventes. Deux, si les propriétaires de l'établissement se fondent sur des estimations, ils peuvent avoir tendance, en déclarant début mars les ventes de janvier, à juger d'après la conjoncture du moment et non pas d'après celle de janvier. Trois, il est très difficile de traiter convenablement les créations et disparitions d'établissements quand une même tournée fournit des données sur deux mois. Par exemple, supposons un établissement en activité début février qui est fermé début mars. Dans le dénombrement en vue d'un sondage aréolaire, l'enquêteur inscrit l'établissement et enregistre ses ventes de janvier au titre de la tournée du cadre de janvier. Toutefois, si cet établissement rentre dans le cadre de février, on oubliera probablement ses ventes de janvier, car il n'existera plus lors de la tournée de ce cadre.

72. Malgré les séduisants avantages du premier type de calendrier, on ne doit pas l'appliquer sans que l'expérience démontre que les déclarations pour le mois en cours et le mois précédent donnent essentiellement les mêmes montants. Il en va particulièrement ainsi parce que l'estimation composite exposée ci-dessous accroît toute différence qui peut exister dans les estimations simples sans distorsion.

73. Pour réduire la variance des estimations de quantité avec les cadres rotatifs, une méthode consiste à utiliser une estimation composite, c'est-à-dire à se servir des déclarations des périodes précédentes pour renforcer l'estimation actuelle. La forme de cette estimation composite dépend du genre de calendrier adopté. Pour le premier genre, cette estimation, que nous appelons Y_i'' (l'indice i désigne le mois) est de la forme :

$$Y_i''' = (1-K)Y_i' + K \frac{Y_i'}{Y_{i-1}''} Y_{i-1}''' \quad (2)$$

Dans cette égalité, supposons encore qu'il s'agit d'une estimation pour janvier. Dans ce cas :

Y_i''' désigne l'estimation composite pour le mois i (janvier);

Y_i' désigne l'estimation actuelle simple et sans distorsion pour le i ème mois (celle pour janvier tirée du cadre de janvier);

Y_{i-1}'' désigne l'estimation simple et sans distorsion pour le mois précédent provenant du mois en cours (c'est-à-dire l'estimation pour décembre tirée du cadre de janvier). Noter que les estimations Y_i' et Y_{i-1}'' proviennent du même cadre

Y_{i-1}''' désigne l'estimation composite pour le mois précédent (décembre) $i-1$ tirée de la même formule;

K désigne une constante qui détermine la pondération respective des deux estimations combinées. Cette constante dépend, à son tour, de la corrélation mensuelle des estimations simples tirées du même cadre. Si nous appelons cette corrélation ρ , la valeur optimale de K est donnée par le paramètre :

$$\frac{1 - \sqrt{1 - \rho^2}}{\rho}$$

Si on fixe K à 0,72 il sera optimal pour $\rho = 0,95$ et quasi optimal quand ρ approche de cette valeur. Avec $\rho = 0,95$, la variance de l'estimation composite (Y_i''') sera à peu près le tiers de celle d'une estimation provenant d'un échantillon non rotatif de même taille.

74. Avec le second type de calendrier, on a deux estimations simples et sans distorsion pour le mois de janvier (l'une provenant du cadre de janvier, que nous appelons Y_i' , et l'autre provenant du cadre de décembre, que nous appelons Y_i''). Noter que chacune de ces estimations se fonde sur des demi-cadres. L'estimation composite en ce cas prend la forme suivante :

$$Y_i''' = K \frac{Y_i'}{Y_{i-1}''} Y_{i-1}''' + LY_i'' + (i-K-L)Y_i' \quad (3)$$

Noter encore que Y_i'' et Y_{i-1}'' proviennent du même cadre (décembre), tandis que Y_i' provient du cadre de janvier. Y_i''' est l'estimation composite calculée selon cette formule pour décembre.

75. Les constantes optimales (fonctions des corrélations mensuelles de cadres identiques) s'écrivent :

$$K = \frac{1 - \sqrt{1 - \rho^2}}{\rho}$$

$$L = \frac{(1 - \sqrt{1 - \rho^2})(1 - \rho)}{2}$$

Si la corrélation mensuelle se fixe à 0,95, les deux constantes prennent la valeur suivante :

$$K = 0,72$$

$$L = 0,04$$

76. L'estimation composite des quantités aura une variance un peu inférieure à la moitié de celle d'un échantillon non rotatif de même taillée.

77. Tandis que les variances des estimations des quantités diminuent sensiblement avec une estimation composite de ce type, celles des estimations des mouvements mensuels empirent. Pour le premier type de calendrier, avec la corrélation et les constantes précédemment admises, la variance s'accroît du tiers environ. Avec le second type, elle est plus que doublée. On doit souligner que, si l'estimation composite nuit aux estimations des mouvements mensuels, on peut l'atténuer grâce aux diminutions souvent très fortes de variance qui s'obtiennent par l'emploi du procédé de grande observation exposé aux paragraphes 84 à 87.

78. Manifestement, l'emploi de cadres rotatifs et de l'estimation composite se justifie bien davantage si on peut obtenir à chaque tournée des renseignements sur deux mois (c'est-à-dire avec le premier type de calendrier). Si ce n'est pas le cas, il se pourrait bien que leur complexité ne soit pas payante.

79. Les variances d'estimation composites (formules 2 et 3) se calculent selon la règle générale du paragraphe 63, c'est-à-dire en calculant pour chaque échantillon tiré dans chaque strate une variable de remplacement U'_{hj} , substituée à Y'_{hj} dans la formule d'estimation (formule 1). Pour les estimations composites provenant de la première formule, cette variable U'_{hj} prend la forme :

$$U'_{(i)hj} = Y'_{(i)hj} + (K) \frac{Y'_{ij}}{Y'_{i-1j}} \left[U'_{(i-1)hj} - Y'_{(i-1)hj} \right] \quad (4)$$

où K est la constante généralisée que nous avons proposé de fixer à 0,72.

80. Les symboles gardent le même sens que dans la formule (2), sauf que l'indice hj indique qu'on ne se sert que des données pondérées provenant du j ième tirage de la h ième strate. Noter toutefois que la

fraction $\frac{Y'_i}{Y'_{i-1}}$ se calcule toujours à partir des totaux globaux.

81. On démontre que la somme $\sum_{h=1}^R \sum_{j=1}^{M_h} U'_{(i)hj}$ est exactement

égale à la valeur de Y'_i calculée selon la formule (2). Donc, on calcule

les variables $U'_{(i)hj}$ pour chaque tirage et on se borne à les additionner pour obtenir les totaux nationaux désirés. On notera aussi qu'il s'agit là d'un procédé en chaîne, en ce sens que les variables U'_{hj} calculées pour le mois i servent avec d'autres données à calculer les variables correspondantes pour le mois $i + 1$.

82. Pour les estimations tirées du second type de calendrier (formule (3)), les variables U'_{hj} prennent la forme suivante :

$$U'_{(i)hj} = (K+L)Y'_{i(hj)} + K \frac{Y'_i}{Y'_{i-1}} (U'_{(i-1)hj} - Y'_{(i-1)hj}) + (i-K-L)Y'_{(i)hj} \quad (5)$$

83. Les symboles gardent le même sens que dans la formule (3), sauf que, de même, l'indice hj indique qu'on ne se sert que des données pondérées provenant du h jième tirage. Les constantes K et L prennent les valeurs 0,04 et 0,72, respectivement, si on applique les approximations empiriques exposées aux paragraphes 73 à 78.

84. Si on se sert d'un ensemble fermé de cadres rotatifs, par exemple sur 12 mois, on peut recourir à une méthode de grande observation pour réduire sans distorsion le poids des aberrances. Cette méthode peut réduire sensiblement la variance, surtout pour le coefficient mensuel. Elle est inécessaire parce que la variance de ce coefficient s'accroît par l'emploi de l'estimation composite, comme il a été dit aux paragraphes 73 à 78.

85. Supposons de nouveau un ensemble de 12 cadres. Le poids de chaque aberrance peut être réduit au douzième de sa valeur initiale pour tout établissement au-dessus de la limite. A cet effet, certes, il faut dénombrer les établissements chaque mois, au lieu de le faire chaque année. Il n'y aura pas de distorsion, pourvu que les établissements soient traités de même quel que soit le cadre où ils se situent. A cet effet, on procède ainsi. Une limite est fixée pour chaque mois de l'année; on l'appelle C_i pour le i ième mois. Cette limite peut être la même pour tous les mois de l'année mais, s'il existe un schéma saisonnier marqué, mieux vaut la faire varier selon ce schéma. La limite est donnée par les valeurs

pondérées (des recettes ou de quelque autre critère). La limite optimale se détermine au mieux par tâtonnements avec les données effectives dès qu'on en dispose (les coûts étant fixes) mais, pour approximation initiale, on peut prendre cinq fois la moyenne pondérée des ventes mensuelles qu'on escompte obtenir du sondage.

86. Les limites ainsi déterminées, on compare les valeurs pondérées pour chaque établissement compris dans un cadre mensuel avec la limite pour ce mois. Si les recettes pondérées dépassent cette limite, l'établissement est placé sur une liste provisoire "de grande observation". En procédant ainsi pour chacun des 12 cadres, on obtient une liste globale qui est vérifiée pour les 11 autres mois. Si les recettes dépassent la limite pour tous, l'établissement est admis dans le cadre "permanent de grande observation" et il fait l'objet d'une enquête dans chaque cadre, son poids étant divisé par 12.

87. Si cette méthode s'emploie conjointement avec l'estimation composite, il n'y a pas lieu de modifier la forme des variables U_{hj} pour calculer la variance. Les valeurs pondérées provenant de ce cadre doivent s'inclure dans toutes les valeurs Y simples et sans distorsion (aussi bien Y' qu' Y'').

88. Avec un sondage aérolaire, il est improbable que la rotation des UPS d'inclusion incertaine soit rapide. Il faut s'attendre qu'elle se borne à celle des ensembles de segments contenus dans ces UPS. En pareil cas, on se sert des mêmes formules et méthodes, mais on n'agit que sur la composante intra-UPS de la variance.

89. Souvent des renseignements auxiliaires, tels qu'emploi, salaires, etc., portés sur des déclarations fiscales récentes, ou des données sur les recettes ou d'autres statistiques obtenues lors d'un recensement récent, peuvent servir à améliorer les estimations courantes. Pour être utiles, ces données doivent être disponibles pour l'univers statistique considéré et on doit pouvoir aussi les évaluer pour l'échantillon qui a été tiré.

90. Pour utiliser ces renseignements auxiliaires, la formule d'estimation suivante est recommandée :

$$Y'_R = \sum_{h=1}^L \frac{Y'_h}{X'_h} X_h \quad (6)$$

où :

Y'_R désigne l'estimation ajustée;

Y'_h désigne l'estimation courante simple et pondérée;

X'_h désigne l'estimation sondée et pondérée de la variable en corrélation (X_h désignant l'univers statistique total de cette variable);

h désigne la hième des L grandes catégories d'activité à l'échelon national. (Noter que ces catégories h peuvent représenter ici des combinaisons des R strates servant au choix).

91. Cette estimation appelle les observations suivantes :

- a) En général, si la corrélation entre X et Y dépasse 0,5, Y'_R aura une variance moindre qu' Y' (estimation simple).
- b) Les L catégories d'activité pour lesquelles on établit des estimations fractionnaires devraient avoir pour X'_h des estimations suffisamment stables (c'est-à-dire avec un coefficient de variation de 15 % ou moins). Les catégories qui ne remplissent pas cette condition doivent être combinées à d'autres.
- c) Si les données que représente le symbole X servent à stratifier et à sonder, l'estimation fractionnaire risque de ne guère apporter d'amélioration supplémentaire.
- d) Si on se sert de cette forme d'estimation fractionnaire, la variable U'_{hj} servant dans la formule concernant la variance prend la forme $U'_{hj} = Y'_{hj} - R'_h X'_{hj}$ où R'_h désigne la fraction $\frac{Y'_h}{X'_h}$ de la formule (6).
- e) Dans un sondage aréolaire, on ne dispose d'ordinaire pas, pour les segments, de données sur les statistiques corrélées. En pareil cas, on prend pour valeur X'_{hj} les totaux pondérés des UPS et cette variable n'influe que sur la composante interstrate de la variance.
- f) Si l'estimation composite sert à réduire la composante intra UPS de la variance dans un sondage aréolaire et si l'estimation par coefficient sert à réduire sa composante inter UPS, l'estimation qui en résulte (appelée $Y'_{(i)R}$) prend la forme suivante :

$$Y'_{(i)R} = \sum_{h=1}^L \frac{Y'_{(i)h}}{X'_h} X_h \quad (7)$$

où :

$Y'_{(i)h}$ désigne l'estimation composite présentée aux paragraphes 73 à 78 pour la h ème grande catégorie d'activité;

X'_h désigne l'estimation pour la h ème grande catégorie d'activité des statistiques auxiliaires tirées des totaux pondérés des UPS pour le h ème genre d'activité;

X_h désigne l'univers statistique total pour ce h ème genre d'activité.

92. Si on se sert de l'estimation ci-dessus, la variable U'_{hj} servant à la formule de la variance (nous l'appelons U''_{hj}) prend la valeur suivante :

$$U''_{hj} = U'_{hj} - R'_h X'_{hj} \quad (8)$$

Dans cette formule

U'_{hj} a la même valeur que dans les formules (4) ou (5), selon le genre de rotation;

R'_h désigne la fraction $\frac{Y''(i)_h}{X'_h}$ appropriée pour le hjième tirage, et

X'_{hj} désigne le total pondéré des UPS pour la hjième grappe finale.

93. Il convient ici de déconseiller la réduction discrétionnaire et massive des poids correspondant aux aberrances dans tout sondage. On est très tenté de le faire en sondant les établissements, car un établissement (ou une entreprise) peut grandir extrêmement entre le moment du sondage sur liste et celui du dénombrement. De même, il est possible que de très grands établissements figurent dans le sondage aréolaire, puisqu'il n'y a pas de stratification (sauf pour une éventuelle liste d'inclusion certaine). Par exemple, supposons qu'un grand magasin figure dans le sondage de la ville A avec un coefficient de pondération égal à 100. Etant donné qu'il s'agit d'une ville moyenne, elle ne peut évidemment pas compter 100 grands magasins. On est donc tenté de réduire discrétionnairement le coefficient (et de le ramener probablement à 1). Cette mesure se justifierait si le seul objet de l'enquête était d'estimer le nombre des grands magasins installés dans la ville A. Toutefois, le total national présente d'ordinaire bien plus d'importance et, à cet échelon, on trouvera bien 100 ou davantage de ces grands établissements. Eliminer le cas considéré ou réduire son coefficient de pondération entraînerait une notable sous-estimation.

94. Il n'existe aucun moyen de dire, d'après l'échantillon, si les cas extrêmes y sont sur-représentés ou sous-représentés. D'ordinaire, les renseignements dont on dispose par ailleurs sont trop limités pour permettre d'en juger à l'échelon global. Donc, les réductions discrétionnaires de pondération doivent s'opérer très sobrement, puisque quelques-unes suffisent à provoquer une forte sous-estimation par distorsion.

95. La méthode de grande observation décrite au paragraphe 86 permet de réduire sans distorsion le poids des aberrances et peut jouer un rôle considérable dans les enquêtes où elles posent un sérieux problème.

Note : L'exposé ci-dessus se consacre spécialement et se limite essentiellement aux techniques pratiques de recours au sondage dans les enquêtes sur la distribution. Il serait bon de lire aussi les documents traitant du sondage qu'énumère la bibliographie, car ils donnent quelques autres méthodes qui pourraient être utiles et exposent certaines questions connexes qui n'ont pas été abordées ici.

Annexe II

BIBLIOGRAPHIE

- Recommandations internationales pour les statistiques de la distribution et des services, série M, No 57, 56 p. Publication des Nations Unies, numéro de vente E.75.XVII.9.
- Classification internationale type, par industrie, de toutes les branches d'activité économique (CITI), série M, No 4, Rev.2. Publication des Nations Unies, numéro de vente E.68.XVII.8.
- Indexes to the International Standard Industrial Classification of All Economic Activities, série M. No 4, Rev.2, Add.1, 247 p. Publication des Nations Unies, numéro de vente E.71.XVII.8.
- Classification type pour le commerce international (deuxième version révisée) (CITI), série M, No 34, Rev.2, 91 p. Publication des Nations Unies, numéro de vente E.75.XVII.6.
- Classification par grandes catégories économiques, série M, No 53, Rev.1, 57 p. Publication des Nations Unies, numéro de vente E.76.XVII.7.
- "Classification internationale type de tous les biens et services" (E/CN.3/493) (Future publication des Nations Unies)
- Bibliography of Industrial and Distributive-trade Statistics, série M, No 36, Rev.4, 177 p. Publication des Nations Unies, numéro de vente E.75.XVII.7.
- Recommandations pour le Programme mondial de statistiques industrielles de 1973, troisième partie. Organisation et exécution des recensements industriels, série M, No 54, 102 p. Publication des Nations Unies, numéro de vente E.72.XVII.10.
- Manuel sommaire des méthodes de sondage, vol. I. Eléments de la théorie des enquêtes par sondage, série F, No 9, Rev.1, 398 p. Publication des Nations Unies, numéro de vente E.72.XVII.5.
- Manuel sommaire des méthodes de sondage, vol. II. Programmes d'ordinateur correspondant aux plans de sondage, série F, No 9, 78 p. Publication des Nations Unies, numéro de vente E.71.XVII.4.
- Manuel sommaire des méthodes de sondage, vol. II. Programme d'ordinateur correspondant aux plans de sondage, série F, No 9, Add.1, 50 p. Publication des Nations Unies, numéro de vente E.73.XVII.8.
- Recommandations pour la préparation des rapports sur les enquêtes par sondage (édition provisoire), série C, No 1, Rev.2, 18 p. Publication des Nations Unies, numéro de vente E.64.XVII.7.
- Enquêtes par sondage d'intérêt actuel (douzième rapport) série C, No 13, 298 p. Publication des Nations Unies, numéro de vente E.73.XVII.5.
- On peut obtenir les publications des Nations Unies à la Section des ventes, Nations Unies, New York ou Genève.