UNITED NATIONS SECRETARIAT **ESA/STAT/AC.93/L6**
Statistics Division **13 November 2003**

English only

Expert Group Meeting to Review the Draft Handbook on Designing of Household Sample Surveys
3-5 December 2003
Venue: Conference Room, 19th floor, DC-2 Building,
2 United Nations Plaza, New York

Outline of the Draft Handbook on Designing of Household Sample Surveys

Introduction.....3

Chapter One: Main sources of socio-demographic statistics 1-4
1.1. Population and housing censuses.....1-4
1.1.1. Basic feature of a population and housing census.....1-4
1.1.2. Uses of census results.....1-5
1.2. Administrative records.....1-5
1.3. Household surveys.....1-6
1.4. Complementarities of the three data sources.....1-8
1.5. Conclusion.....1-9

Chapter Two: Sampling strategies..... 2-3
2.1. Probability sampling versus other sampling methods for household surveys..2-3
2.1.1. Probability sampling.....2-3
2.1.2. Non-probability sampling methods.....2-6
2.2. Sample size determination for household surveys.....2-8
2.2.1. Magnitudes of survey estimates.....2-8
2.2.2. Target population.....2-9
2.2.3. Precision and statistical confidence.....2-9
2.2.4. Analysis groups - domains.....2-10
2.2.5. Clustering effects.....2-12
2.2.6. Adjusting sample size for anticipated non-response.....2-13
2.2.7. Sample size for master samples.....2-13
2.2.8. Estimating change or level.....2-14
2.2.9. Survey budget.....2-14
2.2.10. Sample size calculation.....2-14

2.3. Stratification....2-17
2.3.1. Use of stratified samples....2-17
2.3.2. Rules of stratification....2-18
2.3.3. Implicit stratification....2-19
2.4. Cluster sampling....2-19
2.4.1. Cluster versus stratified sampling....2-21
2.4.2. Cluster design effect....2-21
2.4.3. Cluster size....2-22
2.4.4. Calculating deff....2-23
2.4.5. Number of clusters....2-24
2.5. Sampling in stages....2-24
2.5.1. Benefits of sampling in stages....2-24
2.5.2. Use of dummy stages....2-25
2.5.3. The two-stage design....2-27
2.6. Sampling with probability proportional to size (PPS)....2-28
2.6.1. PPS sampling....2-28
2.6.2. PPES sampling (probability proportionate to estimated size)....2-31
2.7. Options in sampling....2-32
2.7.1. Epsem, pps, fixed-size, fixed-rate sampling....2-33
2.7.2. Demographic and Health Survey (DHS)....2-36
2.7.3. Modified Cluster Design in Multiple Indicator Cluster Surveys (MICS)....2-37
2.8. Special topics – two-phase samples and sampling for trends....2-38
2.8.1. Two-phase sampling....2-38
2.8.2. Sampling to estimate change or trend....2-40

Chapter Three: Sampling frames and master samples.... 3-3
A. Sampling frames and their development....3-3
3.1. Definition of sample frame....3-3
3.1.1. Sample frame and target population....3-3
3.2. Properties of sampling frames....3-4
3.2.1. Completeness....3-4
3.2.2. Accuracy....3-5
3.2.3. Current frame....3-5
3.3. Area frames....3-6
3.4. List frames....3-7
3.5. Multiple frames....3-8
3.5.1. Typical dual frame in household surveys....3-8
3.5.2. Multiple frames for different types of living quarters....3-9
3.6. Typical frame(s) in two-stage designs....3-9
3.7. Master sample frames....3-10
3.8. Common problems of frames and suggested remedies....3-10
3.8.1. Census frame more than two years old....3-11
3.8.2. Census frame two years old or less....3-12
B. Master sampling frames....3-13
3.9. Definition and use of a master sample....3-13
3.10. Ideal characteristics of PSUs for a master sample frame....3-13

3.11. Use of master samples to support surveys 3-14
3.11.1. Advantages of multiple use of a master sample frame 3-15
3.11.2. Disadvantages of multiple use of a master sample frame 3-15
3.12. Allocation across domains (administrative regions, etc.) 3-15
3.13. Maintenance and updating of master samples 3-16
3.14. Rotation of master samples of PSUs 3-17
3.15. Country examples of master samples 3-17

Chapter Four: Documentation and evaluation of sample designs 4-3
4.1. Need for documentation and evaluation 4-3
4.2. Labels for design variables 4-3
4.3. Selection probabilities 4-4
4.4. Response rates and coverage rates at various stages of sample selection 4-4
4.5. Weighting: base weights, non-response and other adjustments 4-5
4.6. Information on costs 4-6
4.7. Evaluation 4-6

Chapter Five: Construction and use of sample weights 5-3
5.1. The need for sampling weights 5-3
5.2. The development of sampling weights 5-3
5.2.1. Adjustments of sample weights for unknown eligibility 5-4
5.2.2. Adjustments of sample weights for duplicates 5-4
5.3. Weighting for unequal probabilities of selection 5-4
5.3.1. A case study in construction of weights: the Viet Nam National Health Survey. 5-6
5.3.2. Self weighting samples 5-6
5.4. The adjustment of sample weights for non-response 5-6
5.4.1. Reducing non-response bias in household surveys 5-7
5.4.2. Compensating for non-response bias 5-7
5.4.3. Non-response adjustment of sample weights 5-8
5.5. The adjustment of sample weights for non-coverage 5-10
5.5.1. Sources of non-coverage in household surveys 5-10
5.5.2. Compensating for non-coverage in household surveys 5-11
5.6. Increase in variance due to weighting 5-12
5.7. Concluding remarks 5-12

Chapter Six: Estimation of sampling errors for complex survey data 6-3
6.1. Introduction 6-3
6.2. The pitfalls of using standard statistical software packages to analyze household survey data 6-3
6.3. Methods of sampling error estimation 6-4
6.4. Preparation of data files for analysis 6-5
6.5. Computer software for sampling error estimation 6-6
6.5.1. CENVAR 6-6
6.5.2. Epi Info 6-6
6.5.3. PC CARP 6-7

6.5.4. SAS 6-7
6.5.5. STATA 6-7
6.5.6. SUDAAN 6-7
6.5.7. WesVarPC 6-7
6.6. Comparison of the software packages 6-7
6.7. Concluding remarks 6-7

Chapter Seven: Nonsampling errors in surveys 7-3
7.1. Introduction 7-3
7.2. Sources of nonsampling errors 7-4
7.3. Types of nonsampling errors 7-5
7.3.1. Specification error 7-5
7.3.2. Coverage or frame error 7-5
7.3.3. Nonresponse 7-7
7.3.4. Measurement error 7-10
7.3.5. Processing errors 7-11
7.3.6. Errors of estimation 7-12
7.4. Bias and variable error 7-12
7.4.1. Variable component 7-13
7.4.2. Systematic error 7-13
7.4.3. Sampling bias 7-13
7.5. Precision and accuracy 7-14
7.6. Assessing on nonsampling errors 7-14
7.6.1. Consistency check 7-14
7.6.2. Sample check/verification 7-15
7.6.3. Post-survey checks 7-15
7.6.4. Quality control techniques 7-16
7.6.5. Study of recall errors 7-16
7.6.6. Interpenetrating sub-sampling 7-17
7.7. Conclusion 7-17

Chapter Eight: Data processing, analysis, and dissemination 8-3
8.1. Introduction 8-3
8.2. A visual summary of the phases and processes of a household survey 8-3
8.3. Survey planning 8-4
8.3.1. Survey objectives 8-4
8.3.2. The tabulation and analysis plan 8-5
8.3.3. Design of the survey questionnaire 8-6
8.3.4. The conceptualization and general considerations for systems design for household surveys 8-7
8.3.5. The options for configuring data processing systems for household surveys 8-12
8.3.6. The development and uses of databases for household surveys 8-16
8.4. Survey operations 8-20
8.4.1. Data collection and data management 8-20
8.4.2. Data preparation 8-20
8.4.3. File structure and datasets for tabulation and analysis 8-29

8.4.4. Estimation and assessment of the accuracy of estimates 8-33
8.4.5. The analysis of survey data 8-38
8.4.6. The reporting, presentation and dissemination of survey data 8-51
Annex 1. Software options for different steps of survey data processing 8-58
Annex 2. Documentation Structure for a survey and its production system 8-59

Annexes..... 3
Annex One: Overview of sample survey design..... A.I.-3
A.I.1. Sample design A.I.-3
A.I.2. Basics of probability sampling strategies A.I.-5
A. Simple random sampling A.I.-5
A.I.2.1. Sample selection A.I.-6
A.I.2.2. Determination of sample size under SRS A.I.-8
A.I.2.3. Coefficient of variation A.I.-10
A.I.2.4. Confidence interval A.I.-10
B. Systematic sampling A.I.-11
A.I.2.5. Linear systematic sampling A.I.-11
A.I.2.6. Circular systematic sampling A.I.-12
A.I.2.7. Estimation in systematic sampling A.I.-12
A.I.2.8. Advantages of using systematic sampling A.I.-13
A.I.2.9. Disadvantages of Systematic Sampling A.I.-13
C. Stratification A.I.-15
A.I.2.10. Advantages of stratified sampling A.I.-15
A.I.2.11. Weights A.I.-16
A.I.2.12. Sample values A.I.-16
A.I.2.13. Proportional allocation A.I.-17
A.I.2.14. Optimum allocation A.I.-17
A.I.2.15. General observations A.I.-18
A.I.2.16. Determination of within stratum sample sizes A.I.-19
D. Cluster sampling A.I.-20
A.I.2.17. Some reasons for using cluster sampling A.I.-20
A.I.2.18. Single stage cluster sampling A.I.-21
A.I.2.19. Sample mean and variance A.I.-21
Annex Two: Planning and execution of surveys A.II.-23
A.II.1. Planning of surveys A.II.-23
A.II.1.1. Objectives of a survey A.II.-23
A.II.1.2. The universe A.II.-23
A.II.1.3. Information to be collected A.II.-24
A.II.1.4. Survey budget A.II.-24
A.II.2. Data collection methods A.II.-25
A.II.3. Questionnaire design A.II.-27
A.II.4. Organization of field work A.II.-29
A.II.4.1. Equipment and Materials A.II.-29