Statistical Training Document
ISP-TR-5

Evaluating Censuses of Population and Housing

U.S. Department of Commerce
Malcolm Baldrige, Secretary
Clarence J. Brown, Deputy Secretary
Sidney Jones, Under Secretary for Economic Affairs

BUREAU OF THE CENSUS
John G. Keane, Director

Issued September 1985

Prepared under a Resources Support Services Agreement
with the U. S. Agency for International Development

ACKNOWLEDGMENTS

The development of this document began in August 1982, when the International Statistical Programs Center (ISPC) of the U.S. Bureau of the Census conducted a regional workshop on methods of census evaluation in Manila, Philippines, attended by participants from nine developing countries: Bangladesh, India, Indonesia, Korea, Nepal, Pakistan, Philippines, Sri Lanka, and Thailand. The training materials and lectures prepared for that workshop provided the basis for portions of this manual.

Evaluating Censuses of Population and Housing was developed by ISPC under the general direction of **Karl K. Kindel,** Chief, ISPC, and **Michael J. Hartz,** Assistant Division Chief for Technical Cooperation, ISPC. Preparation of the manual was coordinated under the direction of **Carol A. Van Horn,** Chief, Census and Survey Methods Branch, and **Nina J. Pane Pinto,** her predecessor. The overall content and organization of the material were under the technical direction of **Anthony G. Turner,** Chief, International Mathematical Statistical Staff, Statistical Methods Division (SMD), and **Robert J. Magnani,** International Mathematical Statistical Staff, SMD.

The chapters constituting this manual can be attributed specifically to the following Bureau of the Census authors: **Robert J. Magnani,** International Mathematical Statistical Staff, SMD, wrote chapters 1, 4, and 5; **Charles D. Cowan,** Assistant Division Chief for Training and Resident Programs, ISPC, wrote chapters 2 and 6; **Paul P. Biemer,** Assistant Division Chief for Mathematics and Statistical Data Analysis, Statistical Research Division (SRD), wrote chapter 3 and the appendix; and **Anthony G. Turner,** Chief, International Mathematical Statistical Staff, SMD, wrote chapter 7.

Several members of the staff of the Center for International Research (CIR), U.S. Bureau of the Census, made important contributions to the demographic materials presented in this document: **Eduardo E. Arriaga,** Special Assistant for International Demographic Methods, and **Arjun L. Adlakha,** Chief, Asia, Europe, North America, and Oceania Branch, provided substantive advice and assisted in the technical review of the materials. **Peter O. Way, Peter D. Johnson, Patricia M. Rowe, Michael K. Roof, Kevin G. Kinsella,** and **Frank B. Hobbs** of CIR assisted in the identification of appropriate illustrative examples and provided technical input. Information on developing country experiences with post-enumeration surveys was gathered under the direction of **Samuel Baum,** Chief, CIR, and **Sylvia D. Quick,** Chief, Latin America and Africa Branch, CIR.

The assistance provided by other persons in the Bureau of the Census and in outside organizations is also gratefully acknowledged. Workshop lecture notes and preliminary manual materials were prepared by **William E. O'Leary** and **David V. Bateman.** Technical review of various portions of the material was conducted by: **H. Bradley Wells,** Professor of Biostatistics at Wake Forest University; **Thomas B. Jabine,** Professorial Lecturer in Statistics at George Washington University; **Charles D. Jones,** Chief, SMD; **Barbara A. Bailar,** Associate Director for Statistical Standards and Methodology; **Howard Hogan,** Chief, Undercount Research Staff, SRD; **Henry F. Woltman,** Chief, Census Data Quality Branch, SMD; **W. Ward Kingkade,** CIR; **Susan Enea, Nancy McGirr,** and **Patricia A. Anderson** of ISPC.

The manual was prepared for printing by the staff of the Census and Survey Methods Branch: **Linda A. Schlueter** coordinated the drafting and preparation of the entire manual. **Dolores E. Adams** and **Carolyn G. Waesche** prepared figures and exhibits. **Barbara McCall** and **Patricia Saukel** typed manuscript portions of the text. The authors also express thanks to the many other persons who directly or indirectly assisted in preparing and improving this publication.

This document was developed under a Resources Support Services Agreement with the U.S. Agency for International Development.

SUGGESTED CITATION

U.S. Bureau of the Census, Evaluating Censuses of Population and Housing,
Statistical Training Document,
ISP-TR-5, Washington, D.C., 1985.

Preface

The purpose of this manual is to describe methods and techniques for the evaluation of censuses of population and housing, with emphasis on techniques which are applicable in developing countries. The manual is intended for use as a basic reference in the design and implementation of census evaluation programs. As such, the rationale and theory behind each method and some of the major problems encountered in its application are covered in sufficient detail so as to guide key decision-makers and technicians of developing country statistical organizations in designing and implementing their own evaluation programs.

In keeping with these objectives, attention in the manual is focused on those methods which are viewed by Bureau of the Census staff as having the widest applicability to developing country situations. Content decisions were based upon the cumulative experience of the Bureau of the Census and statistical organizations in both developed and developing countries which have undertaken census evaluations.

Accordingly, the approaches chosen for emphasis in the manual are: (1) the use of post-enumeration surveys (PES), and (2) the use of demographic methods, particularly those based upon the analysis of two or more successive censuses. Other approaches to census evaluation are also described in the manual but are considered in less detail.

The rationale behind the emphasis of these two approaches is that the post-enumeration survey or PES approach, while being among the more technically and financially demanding of the available approaches, nevertheless may represent the most viable alternative for evaluating census error in countries where data on levels and trends of fertility, mortality, and migration are unreliable or nonexistent. Even where reasonably accurate demographic information is available, the PES approach often provides the only available basis for measuring certain components of census error (age-selective coverage error, for example).

The emphasis on demographic analyses of successive censuses, on the other hand, is based upon two recent developments. First, with the completion of the 1980 round of censuses, most countries now have at least two censuses upon which to base census evaluation efforts. Second, with the participation of numerous developing countries in the World Fertility Survey (WFS) and the United Nations National Household Survey Capability Programme (NHSCP) and continued improvements in the completeness of vital registration systems, many countries are now or soon will be in the position of having sufficiently accurate information on levels and trends in the components of population growth, to make use of these data in undertaking comprehensive demographic analyses of their censuses.

The target audience for this manual includes both higher-level management officials in developing country statistical organizations who are responsible for the overall design and implementation of census evaluation programs, and statisticians and other technical-level personnel who are responsible for conducting evaluation studies. Of primary interest to personnel responsible for planning and managing census evaluation efforts will be the material in chapters 1, 6, and 7. Chapters 2 through 5 and the appendixes at the end of the manual are more technically oriented and are intended to provide guidance to technical-level personnel in the actual implementation of census evaluation studies.

Chapter 1 provides an overview of census evaluation. Chapter 2 describes the use of a PES, or more specifically a post-censal matching study, for measuring census coverage error; it also presents extensions of the basic PES design, alternate design strategies, and examples of PES applications in developing countries. Chapter 3 describes the application of content reinterview studies (conducted either in conjunction with or independently of post-censal matching studies) to the measurement of census content error; it also considers the use of interpenetration studies for measuring different components of content error and developing country experiences in census content error evaluation. Chapter 4 presents the underlying approach, briefly describes the particular demographic techniques which often are useful for census evaluation purposes, and summarizes data requirements and methods for obtaining indirect estimates of the required parameters when reliable direct information is unavailable. Chapter 5 describes these methods in detail and illustrates their application using data from developing countries. Chapter 6 addresses the major issues involved in deciding whether or not and how to adjust census figures on the basis of information obtained from census evaluation studies. Finally, chapter 7 outlines and discusses some of the important factors and considerations involved in planning and implementing a census evaluation effort.

In preparing the manual, an effort was made to present the material in as nontechnical a manner as possible, short of comprising the usefulness of the manual as a basic reference for conducting the types of evaluation studies described. Where feasible, mathematical derivations and other theoretical bases for the methods described are presented separately in the appendixes.

Accordingly, minimal previous background in methods of census evaluation is required for readers to comprehend the logic or general procedures of each method. Previous background in mathematical statistics will, nevertheless, prove useful in considering the material on the design, execution, and analysis of post-enumeration surveys presented in chapters 2 and 3, as will previous background in techniques of demographic analysis in studying the methods presented in chapters 4 and 5. References to appropriate supplementary sources are provided in each chapter to encourage further study of each method.

Finally, while it was intended that the manual be a self-contained reference, it was not always possible for practical reasons to accomplish this objective. For example, it is often necessary for countries without reliable vital registration and immigration statistics to resort to "indirect" methods of estimating fertility, mortality, and migration rates for use in evaluating a census. Because of the large number of these methods which would have to be covered and the fact that they are described in a comprehensive fashion in other widely available sources[1], it was deemed impractical to cover these methods in a systematic fashion in this manual. Accordingly, these supplementary sources are relied upon to provide the necessary background for the estimation of basic demographic parameters. The uses of these estimates for census evaluation purposes are covered in detail in this manual.

[1] See *Manual X: Indirect Techniques of Demographic Estimation,* New York: United Nations, 1983 with regard to fertility and mortality estimation, and Shryock and Siegel, 1975 with regard to estimation of migration.

CONTENTS

Chapter 1. OVERVIEW TO CENSUS EVALUATION

	Page
1. Introduction	1
2. Objectives of evaluation	1
2.1 Products of evaluation	1
2.11 Measures of accuracy	2
2.12 Identification of sources of errors	2
2.2 Uses of the products of evaluation studies	2
2.21 Guide improvements in future censuses and surveys	2
2.22 Assist census data users in interpreting the results	3
2.23 Adjust census results	3
3. Types of census errors	3
3.1 Coverage versus content error	3
3.11 Omissions	4
3.12 Duplications	5
3.13 Erroneous inclusions	5
3.2 Net versus gross error	5
3.3 Sampling versus nonsampling error	7
4. Methodological framework for the evaluation of census errors	8
5. Methods based upon a single source of data	9
5.1 Demographic analysis	10
5.2 Interpenetration studies	11
6. Methods based upon comparisons of data from two or more sources	11
6.1 Matching studies	12
6.11 Post-censal matching surveys	12
6.12 Reinterview surveys	13
6.13 Record checks	14
6.14 Comparison with existing household surveys	15
6.2 Non-matching studies	16
6.21 Demographic analysis using previous censuses	16
6.22 Comparison with administrative records	17
6.23 Comparisons with existing household surveys	17
7. Relative strengths and weaknesses	17
7.1 Methods based upon single versus multiple sources of data	18
7.2 Matching versus non-matching studies	19
7.3 Basis for choice of methodology	21
8. References	22

Chapter 2. STATISTICAL EVALUATION OF COVERAGE ERROR

1. Introduction	23
2. Post enumeration census match study	23
2.1 Dual system estimation	24
2.11 Modeling census misses	24
2.12 The model applied to censuses under perfect conditions	25
2.13 The model applied to censuses under less than perfect conditions	26
2.14 The estimate of the total population size and the net undercount rate	27
2.2 Alternative procedures for misses and erroneous enumerations	27
2.21 Procedure A	27
2.22 Procedure B	28
2.23 Procedure C	28
2.24 Advantages and disadvantages	28
2.3 Alternative definitions for misses and erroneous enumerations	29
2.31 Advantages and disadvantages of Definition I	30
2.32 Advantages and disadvantages of Definition II	31
2.4 Survey and sample alternatives	31
2.41 Specially designed sample for PES *P*-sample	32
2.42 Use of a survey already "in place"	33
2.43 Erroneous enumeration (*E*-sample) design	34
2.5 PES questionnaire design for alternative procedures and definitions	35
2.51 Alternative addresses	36
2.52 Probing questions	36
2.53 Interview control	36
2.54 Census matching information	36
2.6 Matching	36
2.61 Matching procedure	42
2.62 Matching rules	42
2.63 One way versus two way match	48
2.64 Followup interview to obtain additional information	48
2.7 Special problems	48
2.71 Followup of movers	48
2.72 Assignment of current EA codes to the PES	49
2.8 Estimation	49
2.81 Dual system estimation	50

CONTENTS--Continued

Page

2.82 Treatment of missing data...... 51
2.83 Post stratification of estimates.................... 51
2.84 Special problems related to estimation................... 52
2.9 Suggested tabulations............... 52

3. Longitudinal tracing studies............. 53
3.1 Coverage technique.................. 53
3.2 Alternative tracing procedures....... 54

4. Use of network (multiplicity) sampling in estimating coverage error............ 54
4.1 Counting rules...................... 55
4.2 Estimation.......................... 56

5. Profiles of international coverage evaluation studies...................... 56
5.1 Canada.............................. 57
5.2 United States....................... 57
5.21 1950 census post enumeration survey...................... 58
5.22 1960 census post enumeration survey...................... 60
5.23 1980 census post enumeration survey...................... 60
5.3 Korea............................... 60
5.4 Paraguay............................ 60
5.5 Bangladesh.......................... 61
5.6 India............................... 62

6. References............................ 63

Chapter 3. STATISTICAL EVALUATION OF CONTENT ERROR

1. Introduction............................ 65
1.1 The scope of this chapter........... 65
1.2 Sources of nonsampling error......... 65
1.3 The basic concepts of content evaluation......................... 66
1.4 The mean square error of census statistics......................... 67
1.41 *B*, the census bias component.................... 68
1.42 *SRV*, the simple response variance..................... 69
1.43 *CC*, the correlated component of enumerator variance..................... 70
1.44 *MSE(Y)*: putting it all together..................... 72

2. Design of studies to estimate the components of census error.............. 73
2.1 Interpenetration of enumerator assignments to estimate the correlated component............... 74
2.11 Planning....................... 75
2.12 Preparation.................... 76
2.13 Implementation................. 77
2.14 Analysis....................... 77
2.2 The design and implementation of reinterview surveys................ 77
2.21 Sample design.................. 77
2.22 Matching....................... 78
2.23 Timing......................... 78
2.3 Special considerations for reinterview surveys with the simple response variance objective......... 78
2.31 Questionnaire.................. 79
2.32 Interviewers................... 79
2.33 Respondents.................... 79
2.4 Special consideration for reinterview surveys with the response bias objective.................... 79
2.41 Questionnaire.................. 80
2.42 Reconciliation................. 80
2.43 Interviewers................... 81
2.44 Respondents.................... 81
2.45 Processing..................... 81

3. Tabulating, reporting, and interpreting the results.......................... 82
3.1 Estimation and tabulation........... 82
3.11 The estimation of *CC* and ρ_y in enumerator interpenetration studies................ 82
3.12 Estimation of *B* and *I* in reinterview studies.......... 83
3.2 Reporting the results............... 86
3.21 Introduction................... 87
3.22 Data source and definitions... 87
3.23 Field procedures, study design, and estimation methods..................... 92
3.24 Results for housing characteristics................. 92
3.25 Results for population characteristics............. 92
3.26 Data limitations............... 92
3.27 Appendices..................... 94
3.3 Interpretation of the results....... 94

4. Profiles of international content evaluation studies..................... 94
4.1 Bangladesh.......................... 94
4.2 India............................... 95
4.3 United States....................... 96

5. References............................ 98

Chapter 4. OVERVIEW TO CENSUS EVALUATION THROUGH DEMOGRAPHIC ANALYSIS

Page

1. Introduction 99
2. Demographic methods for census evaluation 100
   2.1 Direct comparisons with data from external sources 100
   2.2 Methods based upon comparison with theoretical distributions 103
   2.3 Methods based upon comparison with a previous census or censuses 106
3. Techniques for indirect demographic estimation 109
   3.1 Fertility estimation 110
   3.2 Mortality estimation 112
   3.3 Migration estimation 116
4. References 118

Chapter 5. APPLICATION OF SELECTED DEMOGRAPHIC TECHNIQUES FOR CENSUS EVALUATION

1. Introduction 121
2. Analyses of age-sex distributions 121
   2.1 Graphical analysis 122
   2.2 Age and sex ratios 133
   2.3 Summary indices of error in age-sex data 139
   2.4 Uses and limitations 140
3. Stable population analysis of age distributions 140
   3.1 Basis of method 141
   3.2 Computational procedure 142
      3.21 Step 1: Calculation of the proportional age distribution of the census population 142
      3.22 Step 2: Selection of a model stable age distribution 142
      3.23 Step 3: Comparison of the recorded and stable age distribution 142
   3.3 Examples of stable population theory applications 142
      3.31 1970 census of Thailand 143
      3.32 1960 census of Brazil 146
   3.4 Uses and limitations 149
4. Analyses of successive censuses using actual data on components of population change 152
   4.1 Basis of method 152
   4.2 Data required 154
   4.3 Computational procedure 155
      4.31 Step 1: Adjustment of registered numbers of intercensal births, deaths, and migrants 155
      4.32 Step 2: Computation of the "expected" census population 155
      4.33 Step 3: Calculation of the residual error or error of closure 155
   4.4 Example using population balancing equation 155
   4.5 Uses and limitations 159
5. Analyses of successive censuses using estimates of the components of population change 160
   5.1 Basis of method 160
   5.2 Data required 161
   5.3 Computational procedure 161
      5.31 Step 1: Surviving the population enumerated in the first census forward to the reference date of the second census. 162
      5.32 Step 2: Adjusting for migration 162
      5.33 Step 3: Calculation of the average number of women of childbearing age during the intercensal period 164
      5.34 Step 4: Estimation of the number of births during the projection period 165
      5.35 Step 5: Distributing the estimated number of births by sex 165
      5.36 Step 6: Surviving intercensal births to the end of the projection period 165
      5.37 Step 7: Comparison of the expected and enumerated census populations 166
   5.4 Examples 166
      5.41 1980 census of the Philippines 166
      5.42 1980 census of Indonesia 172
   5.5 Uses and limitations 176
6. Analysis of cohort survival rates 178
   6.1 Basis of method 179
   6.2 Data required 179
   6.3 Computational procedure 179
      6.31 Step 1: Adjustment for migration 180
      6.32 Step 2: Calculation of census cohort survival rates 180
      6.33 Step 3: Calculation of life table survival rates 180
      6.34 Step 4: Calculation of cohort survival ratios 180
   6.4 Examples 180
      6.41 1960 to 1980 censuses of the Philippines 180
      6.42 1881 to 1961 censuses of India 185
   6.5 Uses and limitations 186
7. Two-census regression method 187
   7.1 Basis of method 187
   7.2 Data required 189
   7.3 Computational procedure 189
      7.31 Step 1: Calculation of the population aged x years and above for the two censuses 190

CONTENTS--Continued

Page

7.32 Step 2: Calculation of the number of intercensal deaths........................ 190
7.33 Step 3: Computation of cumulated proportions dead and census survival ratios.... 191
7.34 Step 4: Calculation of regression parameters......... 191
7.4 Example............................ 191
7.5 Uses and limitations................ 195
8. References............................. 195

Chapter 6

Chapter 6. ADJUSTING CENSUS FIGURES

1. Introduction.............................. 199
2. General considerations in deciding to make an adjustment...................... 199
   2.1 The initial decision to make an adjustment........................ 199
   2.2 Specific issues in considering adjustment........................ 201
3. Impact of varied uses of census data...... 202
4. Operational considerations in making an adjustment............................ 203
   4.1 Timing and revisions................ 204
   4.2 Integrating the adjustment into the census process...................... 204
5. Methodological considerations in making an adjustment.......................... 205
   5.1 Measurement techniques............... 205
   5.2 Methods of adjustment................ 205
   5.3 Implementation of methods............ 207
6. Evaluation of the evaluation............. 208
7. Conclusions.............................. 208
8. References............................... 209

Chapter 7. PLANNING FOR CENSUS EVALUATION

1. Introduction.............................. 211
2. General requirements...................... 211
   2.1 Overall objectives................... 211
      2.11 Scope......................... 212
      2.12 Cost.......................... 212
      2.13 Resources..................... 213
      2.14 Feasibility................... 213
   2.2 Organization, recruitment, and training.......................... 213
   2.3 Preparation of overall technical plan.............................. 214
      2.31 Relation to overall objectives. 214
      2.32 Coordination with census operations.................... 214
      2.33 Manageability................. 215
      2.34 Choice of methodologies........ 215
      2.35 Development of design.......... 215
   2.4 Budgeting........................... 215
   2.5 Development and monitoring of schedules........................... 216
   2.6 Publicity, confidentiality, and legal mandate...................... 216
3. Special requirements for field studies................................. 217
   3.1 Questionnaire design and survey plan............................... 217
   3.2 Sample design....................... 217
      3.21 Domains....................... 217
      3.22 Sampling frame................ 217
      3.23 Independence.................. 218
   3.3 Mapping............................. 218
   3.4 Pretesting.......................... 218
   3.5 Matching............................ 218
   3.6 Preparation of estimates and measures of precision............... 219
4. Special requirements of demographic techniques............................. 219
   4.1 Staff development................... 219
   4.2 Overall measurement methodology...... 219
   4.3 Obtaining independent data sources............................. 220
5. Coordinated planning................. 220

APPENDIX

1. Introduction............................ 223
2. The general model and formulation......... 223
3. Error structures and their use............ 227
   3.1 Quantitative data.................... 227
   3.2 Qualitative data..................... 229
   3.3 Error structure summary.............. 233
4. The estimation of response error components............................. 233
   4.1 Interpenetration study methodology... 233
      4.11 Full interpenetration.......... 234
      4.12 Two-staged interpenetration.... 236
      4.13 Precision...................... 239
   4.2 Reinterview surveys.................. 240

CONTENTS--Continued

	Page
4.21 Notation for quantitative data	241
4.22 Notation for qualitative data	242
4.23 Estimation for both types of data	243
4.24 Covariance between trials	245
4.25 Precision of the estimators	245
5. References	247

LIST OF FIGURES

1-1	Diagram showing various types of census errors	7
1-2	Typology of methods for the evaluation of census errors	9
1-3	Developing countries in which post-enumeration surveys were conducted in conjunction with the latest census	21
2-1	Distribution of captures in a capture-recapture study	25
2-2	Comparison of definitions I and II	30
2-3	Illustrative post enumeration survey questionnaire	37
2-4	Illustrative erroneous enumeration survey questionnaire	43
2-5	Comparison of PES estimates of 1950 net undercoverage with demographic analysis estimates	59
2-6	Comparison of census evaluation study (PES) estimates of 1960 net under-coverage with demographic analysis (DA) estimates of 1960 net undercount	59
2-7	1981 Bangladesh census coverage estimates and their standard errors	62
2-8	1981 India census coverage estimates and their percentage relative standard error	63
3-1	Census and birth record classification of 1,000 census respondents by age	69
3-2	Proportion of persons 14 years of age or less by enumerator assignment	71
3-3	Equivalent effects of enumerator variance and simple response variance, by level of bias, for larger populations	73
3-4	Equivalent effects of enumerator variance and simple response variance, by level of bias, for larger populations	73
3-5	1980 U.S. census questionnaire	80
3-6	1980 U.S. census content reinterview questionnaire	80
3-7	Notation for computing the correlated component estimator and its standard error	83
3-8	Computing $\widehat{CC}$ and $\hat{\rho}_y$	84
3-9	Computing the standard errors and ninety-five percent confidence intervals	85
3-10	Example of procedure for three interpenetrated enumerator pairs	86
3-11	General notation for computing response error measures in reinterview studies	87
3-12	Example of procedure for computing response error measures	87
3-13	Computing net difference rate and index of inconsistency	88
3-14	Computing standard errors and ninety-five percent confidence intervals	89
3-15	Example of a detailed reinterview table	93
3-16	Example of a summary table for reporting measures from reinterview studies	93
3-17	Example of summary table for reporting measures of enumerator variance	94
3-18	General guidelines for interpreting the magnitudes of response error measures for censuses	95
3-19	Aggregate index of inconsistency for selected characteristics: Bangladesh.	95
3-20	Response variance and response bias by sex, literacy, and age	96
3-21	Values of ρ_y for selected characteristics, U.S. population: 1950	97
4-1	Female marital distributions from the 1971 census and 1972 fertility survey: All-India	103
5-1	Census population of Japan by age and sex, 1950	123
5-2	Census population of Japan by age and sex, 1980	123
5-3	De facto census population of Yemen by age and sex, 1975	125
5-4	Census population of Argentina by age and sex, 1980	125
5-5	De facto population of Yemen by age and sex, 1975	125
5-6	Census population of Argentina by age and sex, 1980	127
5-7	De facto census population of Yemen by sex and single years of age, 1975	128
5-8	Census population of Argentina by sex and single years of age, 1980	129
5-9	De facto census population of Yemen, males by single years of age, 1975	131
5-10	Census population of Argentina, males by single years of age, 1980	131
5-11	Graphical cohort of census enumerations of males in the Turkish censuses of 1970, 1975, and 1980	132
5-12	Age ratios for five-year age categories by sex: 1975 census of Yemen and 1980 census of Argentina	134
5-13	Census of population by age and sex and sex ratios: 1975 census of Yemen and 1980 census of Argentina	135
5-14	Sex ratios of the census populations in urban and rural areas in Argentina, 1980	136
5-15	Sex ratios of the population enumerated in the most recent censuses of selected countries	137

CONTENTS--Continued

		Page
5-16	Sex ratios by age group for the population enumerated in the 1951, 1961, 1972 and 1981 censuses of Pakistan.....	138
5-17	Sex ratios by age group for the population enumerated in the 1951, 1961, 1972, and 1981 censuses of Pakistan....	138
5-18	Sex ratios of birth cohorts as enumerated in successive censuses in Pakistan, 1951 to 1981..................	139
5-19	Population enumerated in the 1970 census of Thailand, by age and sex.....	143
5-20	Comparison of age distribution of the population enumerated in the 1970 census of Thailand and a stable age distribution, by sex..................	144
5-21	Ratio of the proportion of the population enumerated at each age in the 1970 census of Thailand to a stable age distribution......................	145
5-22	Age distribution and fitted birth-rate estimates for males, Brazil, 1960......	147
5-23	Plot of the estimated birth rate, $b(x)$, sequence based on the model stable male population, west family, level 13, Brazil, 1960......................	147
5-24	Birth rate estimates, $b(x)$, obtained from proportions under age x $C(x)$, in model stable populations for selected African female populations............	149
5-25	Birth rate estimates, $b(x)$, obtained from proportions under age x, $C(x)$, in model stable populations for selected Latin American male populations and Philippine males............	150
5-26	Stable population analysis of the reported and adjusted age distribution: Paraguay 1972........................	151
5-27	Registered and adjusted numbers of births, deaths, and net migrants in Sri Lanka during the 1971-1981 intercensal period........................	156
5-28	Movement of passengers into and out of Sri Lanka, 1976 to 1981...............	158
5-29	Enumerated 1970 and 1980 census population of the Philippines, by age and sex................................	166
5-30	Estimated and projected infant mortality rates and life expectancy at birth for the Philippines, selected years during the 1970-1980 intercensal period..............................	167
5-31	Values of ${}_5L_x$ and ${}_5S_x$ from life tables for the 1970-1975 and 1975-1980 periods by age and sex, Philippines....	168
5-32	Worksheet for applying the cohort-component methods to the 1980 census of the Philippines...................	169
5-33	Estimated number of net emigrants from the Philippines by age and sex: 1970-1975 and 1975-1980..............	170
5-34	Estimated and projected age-specific and total fertility rates for the Philippines for selected years, 1968 to 1980..............................	170
5-35	Worksheet showing computations of the estimated number of annual births in the Philippines: 1970-1975 and 1975-1980........................	171
5-36	Enumerated and expected 1980 census population of the Philippines, by age and sex.........................	173
5-37	Enumerated and adjusted 1971 census population and enumerated 1980 census population of Indonesia, by age and sex................................	174
5-38	Estimated and projected infant mortality rates and life expectancy at birth for Indonesia, selected years during the 1971-1980 intercensal period......	175
5-39	Values of ${}_nL_x$ from life tables corresponding to estimated and projected levels of mortality for Indonesia, selected years during the 1971-1980 intercensal period..................	176
5-40	Comparison of the enumerated and expected 1980 Indonesian census populations, by age and sex..............	177
5-41	Estimated and projected age-specific and total fertility rates for Indonesia for selected years during the 1971-1980 intercensal period......	178
5-42	Cohort survival rates for selected cohorts in the Philippines for the 1960-70, 1970-75, and 1975-80 intercensal periods.......................	181
5-43	Life table ${}_nL_x$ values and survival probabilities for intercensal periods in the Philippines, 1960-1980.........	182
5-44	Observed female intercensal cohort survival rates and estimated life-table survival rates for the 1960-70, 1970-75, and 1975-80 intercensal periods in the Philippines...........	183
5-45	Estimated number of net female emigrants from the Philippines during the 1975-1980 intercensal period, by cohort................................	185
5-46	Unadjusted and adjusted female intercensal cohort survival and life-table survival rates for the Philippines, 1975-1980 intercensal period..........	185
5-47	Census survival ratios for selected cohorts in Indian censuses from 1881-1961................................	186
5-48	Lexis diagram illustrating the major concepts of the two-census regression method................................	188
5-49	Smoothed populations generated in the 1961 and 1971 censuses of India, by 5-year age groups....................	191
5-50	Smoothed cumulated populations from the 1961 and 1971 censuses of India, estimates of cumulative number of intercensal deaths, and estimated ratios $(D^*/P_0$ and $P_1/P_0)$.............................	192

Page

5-51 Life table ${}_5L_x$ and T_x values and 10-year nonsurvival probabilities for India for the 1961-1971 intercensal period....... 193
5-52 Plot of points $(D^*/P_0, P_1/P_0)$ derived from the smoothed age distribution from the 1961 and 1971 censuses of India................................ 194
5-53 Estimates of the level of coverage in the 1971 census of India relative to the 1961 census.................... 195

a. Theoretical values of I for compensating false positives and false negatives............................ 231
b. Theoretical values of I for noncompensating false positives and false negatives....................... 231

Page

c. Values of ρ_y and ρ_y^* for given P, θ, σ_θ^2, ϕ, and σ_ϕ^2....................... 232
d. Summary of nonsampling error structures.......................... 234
e. Analysis of variance table for computing $\hat{CC}$......................... 235
f. Estimation of CC...................... 237
g. Two-staged interpenetration design for $k = 8$, $a = 2$, and $\ell = 2$........... 238
h. Cross-classification of original interview results by reinterview results.............................. 242
i. Probabilities for the original interview and reinterview cross-classification...................... 242
j. Distribution of census and reinterview responses for vocational training..................... 246

GLOSSARY

Terms and definitions....................... 251

Chapter 1. OVERVIEW TO CENSUS EVALUATION

1. INTRODUCTION

In recent decades substantial improvements in the taking of censuses of population and housing have been realized in many countries. This is due in some measure to an increasing awareness of the existence of errors in census data on the part of both producers and users of the information. While the initial effect of this awareness has been to destroy faith in the absolute accuracy of census results, the more lasting result is likely to be a sounder and more defensible view of census-taking. Here census evaluation plays a vital role.

A "perfect" census is impossible; errors inevitably occur. Nevertheless, census figures that are subject to error are still valuable if the limitations of the data are understood by the users and if the errors do not adversely affect the major uses of the data. Few decisions are likely to depend on knowledge that a country's population is exactly 21,728,516 persons; but decisions may well depend on the determination that the population is between 21.0 million and 22.5 million.

The stage has now been reached in the field of survey sampling that sampling error is readily measurable and controllable, to the extent that sampling errors are probably less problematic relative to other types of error (nonsampling errors) which affect survey results. For censuses in which the population is enumerated on a 100 percent basis, there is no sampling error; however, just as in survey operations, census personnel introduce nonsampling errors. Some assessment of the magnitude and direction of these errors is necessary to respond to questions about the results and attacks on their accuracy.

Accordingly, evaluation studies which examine the results of and the procedures and operations used in undertaking a census are necessary to provide both the producers and users of the data with information needed to assess census quality. Such studies provide users with a basis for deciding either that the errors are relatively small and not likely to affect most conclusions drawn from the data or that the errors are relatively large and inferences should be made with caution.

2. OBJECTIVES OF EVALUATION

In considering evaluation objectives, an initial distinction should be made between the products of evaluation programs and the ways in which those products are used. In the first category are various measures of the accuracy of census data, and information about sources of error. The products of evaluation efforts can be used in several ways: (a) to guide improvements in future censuses and surveys, (b) to assist census data users in their interpretation of the results, and (c) as a basis for adjusting the census results.

2.1 *Products of evaluation*

The products of evaluation are measures of census error and identification of the sources of this error.

2.11 Measures of accuracy.--Within the budgetary constraints of a census program, the first priority of an evaluation effort is to measure the accuracy of at least some of the key census statistics. Accuracy represents the quality of a census result and is measured by the difference between the census figure and the true value of the characteristic being measured. While the true value is seldom known, it can often be approximated.

When demographic methods are used to evaluate a census, accuracy is measured by the net census error, which is the difference between what is considered to be the correct figure and the actual census figure.

In a more statistical (i.e., stochastic) approach to census evaluation, accuracy is measured by trying to estimate the relative importance of various components of census error. Error in a census statistic can arise at any stage in the census process from such sources as varying interpretations of questions by enumerators and/or respondents, unwillingness or inability to give correct answers, nonresponse, coding errors, and other processing errors exclusive of sampling error. Some of these errors are systematic in nature and will not cancel each other out, giving rise to biases. Others tend to be random in nature and balance out (theoretically) over repeated trials or over a large number of interviewers, coders, supervisors, etc., giving rise to variance.

2.12 Identification of sources of errors.--Another primary objective of evaluation is to identify the major sources of error so that future census operations can be conducted more accurately and/or cost effectively than the present census.

The following are illustrative of the types of questions for which evaluation exercises may provide information. At what stage was the error introduced? Was it caused by faulty listings, a poorly designed questionnaire, ill-trained enumerators, subjective coding operations, computer programmer mistakes, or at another stage? Does the error affect one section of the population more than another?

2.2 *Uses of the products of evaluation studies*

The outputs of the evaluation program are utilized in several ways. These include guiding improvements in future censuses and surveys, assisting census data users in interpreting the results, and adjusting census results.

2.21 Guide improvements in future censuses and surveys.--The evaluation program provides both the producer and the user of the data with valuable information for planning future censuses in order to meet data needs more adequately. For example, it may be that a particular geographic or demographic group of national importance to decision-makers was underenumerated. As a result, it may be concluded that new techniques for enumerating them more accurately will need to be developed for the next census. Or, it may be concluded that other methods of questionning on particular topics are likely to produce more accurate results.

The objective of any census should be to achieve the desired degree of accuracy of results for the lowest cost. When choosing among alternative methods of data collection, the relative levels of the nonsampling error for each method need to be known in order to determine the most cost-effective method. Evaluation studies can provide information on the relative performance of various methods to aid in this decision-making process.

2.22 Assist census data users in interpreting the results.--Since the producers of census data tend to be the most knowledgeable regarding the procedures used in collecting and processing the data, they are in an ideal position to provide guidance to data users with regard to the limitations of the data. The dissemination of the results of census evaluation studies serves two useful purposes: (a) it alerts users to the fact that there are errors present in the data, and (b) it provides information on the relative magnitude of error for particular data items and possibly on the relative importance of various sources of error. This information can then be used in determining the proper degree of confidence to be assigned to conclusions and inferences derived from the data, as well as in making necessary adjustments for particular purposes.

2.23 Adjust census results.--As noted above, census evaluation studies provide a basis for assessing the need to adjust the census data to compensate for the effects of errors in the statistics, as well as useful information regarding the nature and magnitude of the required adjustments.

It should be noted, however, that decisions on the necessity and the methods to be used to adjust census results involve a number of important and potentially sensitive considerations. Chapter 6 discusses the feasibility and some of the problems and means of adjusting census data to conform more closely with what is thought to be actual population parameters.

3. *TYPES OF CENSUS ERRORS*

Census errors can arise from several sources such as less-than-perfect data collection and processing procedures or poor operational control resulting in the loss of documents or erroneous coding and keying. Census designers also contribute to errors by producing faulty instruments, instructions, training materials, and procedures: Errors are intrinsic to the nature of a large scale data collection effort such as a census in spite of efforts to avoid them.

In view of the numerous ways in which errors can enter into a census operation, a useful starting point in thinking about methods to measure them is to organize the various types of errors into analytical categories. For census evaluation purposes, classifying errors in terms of the following dimension provides a useful analytic framework: (a) coverage versus content error, (b) net versus gross error, and (c) sampling versus nonsampling error.

3.1 *Coverage versus content error*

Perhaps the most fundamental distinction to be made between types of census errors is between errors of coverage and errors of content. Coverage error is the error in the count of persons or housing units resulting from cases having been "missed" during census enumeration or counted erroneously either through duplication or erroneous inclusion. Content error, on the other hand, is defined as error in the recorded characteristics of those persons that were enumerated in the census. Both coverage and content errors affect the distribution of the population recorded in the census with respect to census characteristics.

Coverage error arises in census enumerations due to such factors as defective field operations, carelessness on the part of enumerators, misunderstanding or lack of cooperation on the part of respondents, or simply because census forms are lost or destroyed during the census processing operation. Content error in population and housing statistics can result from such things as erroneous or inconsistent reporting of characteristics by respondents, failure on the part of enumerators to obtain or record accurately the required information, errors introduced in the clerical and processing operation, etc.

An important point to be made in connection with coverage and content error is that it is of critical importance for evaluation purposes that the concepts and characteristics being measured in the census be clearly defined. For example, one cannot measure census coverage error unless there is a clear and unambiguous definition of the target population for the census. At the most basic level, this involves a choice between a counting rule based upon "usual" residence (a *de jure* counting rule) or one based upon actual residence at the time of the census (a *de facto* counting rule). Other critical decisions concern how aliens residing in the country and nationals residing abroad are to be handled and the definitions of key concepts such as "usual", "temporary", "residence", "household", etc. Similarly, content error cannot be measured in any meaningful way unless there are clear and uniformly applied definitions of key census characteristics such as age, marital status, and income.

Within the category of coverage error, it is important to distinguish between three types of errors in coverage: (a) omissions, (b) duplications, and (c) erroneous inclusions. The reason for this distinction is that it is desirable in post-censal matching studies of census coverage (commonly referred to as post-enumeration surveys - see chapter 2) to obtain separate estimates of each of three types of coverage error so that the overall (or net) coverage error can be estimated.

3.11 Omissions.--Omissions result from (a) entire housing units, households, or persons having no established place of residence (nomads, for example) having been missed by census enumerators or (b) from one or more persons within enumerated housing units or households being missed. In the case where an entire housing unit is missed, it follows that all households and persons residing within the housing unit will also be missed by the census.

There are two primary causes of omission of housing units: (a) failure to include part of the land area of the country in creating enumerator assignments, and (b) enumerator canvassing error within assigned areas. The former problem can be caused by such factors as imprecise boundaries of geographic or census administrative units, faulty maps, or simply by coverage errors made by field staff in the pre-census listing operation. Enumerator canvassing errors can result from such factors as imprecise definition of enumerator assignments, faulty maps, or simply oversight on the part of the enumerators. These errors tend to occur more frequently in sparsely settled rural areas where villages are separated by large distances and in densely settled urban areas characterized by multi-unit structures which are arranged in a complex fashion.

In addition to having been missed because the housing unit has been missed, households can be missed because all of the members of the household were at another place of residence at the time of census enumeration, were temporarily absent during the hours of census enumeration (working, at school, etc.), or were in transit either within or outside of the country during the enumeration period. The likelihood of omission of households tends to be higher in situations where the presence of more than one household in a housing unit is more common than when there is a one-to-one correspondence between housing units and households.

Omissions of persons within enumerated households also occur for a variety of reasons, including deliberate or inadvertent omission of household members on the part of household respondents or application of an incorrect definition of the household by census enumerators. Very young children and adults aged 15-29 tend to be over-represented among persons missed in census enumerations based upon experience to date in a wide variety of countries. It is thought that children tend to be underenumerated most frequently due to respondent recall error, while persons aged 15-29 are in the age range in which the rate of residential mobility is high in many countries and thus these persons tend to be omitted due to uncertainty on the part of respondents and enumerators as to where they should be enumerated.

3.12 Duplications.--Duplications occur when housing units, households, or persons are enumerated more than once in a census. Frequently, duplication is caused by the "overlapping" of enumerator assignments due to errors made during pre-census listing or enumeration area (EA) delineation, or because of the inability of enumerators to identify on the ground the proper boundaries of a particular enumerator assignment. Persons who are mobile residentially or have more than one residence are especially prone to being enumerated more than once.

As a result of duplications, it is conceivable that a census count of the total population may be larger than the actual population. It is more frequently the case, however, that the number of omissions exceeds the number of duplications, resulting in a net census undercount. (See Section 3.2 of this chapter for a discussion of net census error).

3.13 Erroneous inclusions.--These include housing units, households, and persons that were enumerated in the census and either should not have been or were enumerated in the wrong place. Examples of erroneous inclusions are persons who died before the census, who were born after census day, or aliens. A real problem has been noted in some countries of "fictionalized" housing units, households, or persons. This has been especially true in cases where enumerators are paid on the basis of the number of units, households, or persons that they enumerate. Also, persons can be enumerated in the wrong geographic area, resulting in overenumeration for that area and undercoverage for the area where they should have been enumerated. Erroneous inclusion of aliens sometimes occurs in countries whose census is de jure, as opposed to de facto. (See the glossary at the end of this manual for the definition of de jure and de facto). Less ambiguity arises in a de facto census, in which everyone actually residing in the country including aliens is counted, whereas in a de jure census only those persons who "usually" reside in the country are counted.

3.2 *Net versus gross error*

A second important distinction for census evaluation purposes is between net error and gross error. Gross error refers to the total number of errors made in the census, while net error refers to the total effect of these errors on the resultant statistics.

Gross errors affect the nonsampling variances, while net errors affect the nonsampling biases.

In the case of coverage error, for example, gross census coverage error would consist of the total of all persons omitted plus all duplicates plus all erroneous enumerations. In measuring net census coverage error, however, the fact that one of these types of error results in an underestimate of total population (omissions) while the other two types (duplications and erroneous enumerations) result in overestimates is taken into account. Thus, net census coverage error would be measured by the excess or deficit of errors resulting in population underestimates over those errors resulting from population overestimates. A net census undercount is said to exist when the number of omissions exceeds the sum of the number of duplications and the number of persons erroneously enumerated, while a net census overcount is said to exist when the number of duplications plus erroneous enumerations exceeds the number of omissions.

To illustrate the concepts of gross and net content error, consider an age distribution of persons enumerated in a census. Measures of gross content error would consider all cases in which an age other than a respondent's actual age was recorded in the census as errors, while measures of net content error would consider only those errors which are not cancelled out or compensated for by other errors. For example, if N persons whose actual age was X years reported their age as something other than age X in the census, but an equal number of persons (N) whose actual age was something other than X years reported their age as X, there would be no net content error for the census count for persons aged X years, since the total count of persons aged X years would be correct. Each of these types of errors would, however, be counted as errors by measures of gross error.

In actual practice, the measurement of net and gross census error is somewhat more involved and based upon more sophisticated statistical methods than the simple illustrations presented here. These simple illustrations are intended to convey the basic ideas involved.

A more systematic treatment of the relationships between coverage/content and net/gross errors is provided by the following example:

Let P_x^t = the actual or true count of persons in category x in the population (in practice, this is unobservable)

P_x^c = the census count of persons in category x

a = persons in category x correctly included in the census and correctly reported in category x

b = persons in category x incorrectly omitted from the census

c = persons in category x correctly included in the census but reported in a category other than x

d = persons incorrectly included in the census and reported in category x

e = persons correctly included in the census but incorrectly reported in category x

In this example, the correct population count for category x can be expressed, in terms of figure 1-1, as:

$$P_x^t = a + b + c$$

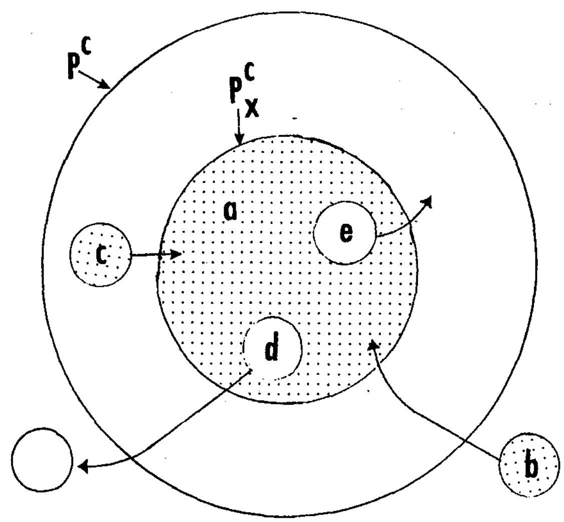
The census count can be expressed as:

$$P_x^c = a + d + e$$

Thus, the net census error for category x may be expressed as:

$$P_x^t - P_x^c = a + b + c - (a + d + e)$$
$$= b + c - d - e$$

Figure 1-1. DIAGRAM SHOWING VARIOUS TYPES OF CENSUS ERRORS



where: P^c = the population counted in the census

P_x^c = population reported in category x

NOTE: Shaded areas indicate groups truly belonging to category x, and the arrows with attached circles indicate where the groups b, c, d, or e <u>should have</u> been counted but were not.

The result of $b + c - d - e$ may be negative or positive. It represents the net error for the census category. Other estimates for a category x are defined as follows:

$b + c$ = census gross undercount

$d + e$ = census gross overcount

Sources of error:

$b + d$ = gross error due to coverage problems

$b - d$ = net error due to coverage problems

$c + e$ = gross error due to content misclassification

$c - e$ = net error due to content misclassification

Thus, error associated with "b" is not due to misclassification. Group "b" is in the correct category, but was incorrectly omitted from the census. Group "d" was erroneously enumerated and possibly misclassified. These are both problems of coverage. Groups "c" and "e" were not missed from the census, but were misclassified and thus are content errors.

3.3 *Sampling versus nonsampling error*

A final important distinction to be made in considering types of errors encountered in population and housing censuses is between sampling and nonsampling errors. Sampling error arises because information is not collected from the entire target population, but rather only from some portion of the population. Through the use of scientific sampling procedures, however, it is possible to estimate the range within which the true population value or parameter is likely to be with a known probability from the sample data.

Nonsampling error, on the other hand, is defined as a residual category consisting of all other errors which are not the result of the data having been collected from only a sample rather than the entire target population. These include errors made by respondents, enumerators, supervisors, office clerical staff, key punch operators, etc.

Experience suggests that for many, if not most, sample survey efforts in both developing and developed countries the contribution of nonsampling error to total survey error then exceeds that of sampling error. Nonsampling errors are likely to have a major impact in a large-scale data collection program such as a population and housing census.

Of course, in a census in which the population is enumerated on a 100 percent basis for all data items, there is no sampling error. In this case, a census evaluation program would be directed entirely toward the measurement of nonsampling error. For censuses in which a subset of census data items are measured on a sample basis, both sampling and nonsampling error would normally be assessed as part of the census evaluation program.

Discussion of sampling error (that is, errors encountered in the census sampling operation, consisting of both sampling variance and sampling bias) is not considered in detail in this manual for two primary reasons: (a) the use of sampling in censuses does not pose any unique problems beyond those encountered in any large-scale survey undertaking, and (b) there are numerous authoritative references on applications of sampling theory available elsewhere such that a comprehensive treatment in this manual would be redundant (Kish 1965; Hansen et al. 1953; Cochran 1953).

Accordingly, the primary focus of the material presented in this manual concerns the evaluation of nonsampling error in censuses of population and housing. Some sampling issues are covered in connection with the design of post-enumeration surveys in chapters 2 and 3.

4. *METHODOLOGICAL FRAMEWORK FOR THE EVALUATION OF CENSUS ERRORS*

There exists a fairly large number of methods which can be applied in census evaluation situations. While the methods differ widely in terms of level of technical sophistication, data requirements, and quality of results, a useful analytic framework or typology can be created by grouping methods on the basis of three criteria: (a) whether a single source of data (the census itself) or more than one source of data is needed, (b) for methods requiring multiple sources of data, whether or not matching on a record-by-record basis is required, and (c) the type of error to be measured (coverage or content, gross or net). A typology of census evaluation methods based upon these three criteria is shown in figure 1-2.

As indicated in figure 1-2, the available options for census evaluation purposes are quite limited when only the results of the census being evaluated are available. In such a case, only a handful of demographic methods are available. These methods are discussed in section 5 of this chapter and described in more detail in chapters 4 and 5. If work assignments for the census are arranged in a manner such as that described in chapter 4, the additional option of evaluating the effects of operator variance on the census results becomes available.

The existence of additional sources of data other than the census itself opens up a much wider range of options for census evaluation purposes. Under the heading of matching studies (that is, studies in which census records are matched on a one-to-one basis with records from another data source) are (a) post-censal matching surveys, (b) re-interview surveys, (c) administrative record checks, and (d) comparison with existing household surveys. These methods are described briefly in section 6 of this chapter and in a more comprehensive fashion in chapters 2 and 3.

Under the heading of non-matching studies are (a) demographic analysis using previous censuses, (b) comparison with administrative statistics, (c) comparison with data from existing household surveys, and (d) interpenetration studies. Brief descriptions of these methods may be found in section 6 of this chapter and more systematic treatments in chapters 4 and 5.

Figure 1-2. TYPOLOGY OF METHODS FOR THE EVALUATION OF CENSUS ERRORS

Source(s) of Data and Methods	Type of Error			
	Coverage Error		Content Error	
	Net	Gross	Net	Gross
<u>Single Source of Data:</u>				
Demographic analysis of the census	x[1]		x[1]	
Interpenetration studies conducted as part of the census		x[2]		x[2]
<u>Multiple Sources of Data:</u>				
Matching studies,				
Post-censal matching surveys	x[3]	x[3]		
Reinterview surveys			x[3]	x[3]
Record checks	x	x	x	x
Comparison with existing household surveys	x	x	x	x
Non-matching studies,				
Demographic analysis using previous censuses	x[1]		x[1]	
Comparison with administrative statistics	x[1]		x[1]	
Comparison with existing household surveys	x[1]		x[1]	

[1]*As a practical matter, these methods do not enable the analyst to evaluate the relative contributions of coverage and content to total error. Useful information can, however, be obtained through the use of demographic models and the comparison of successive censuses (see Chapter 5).*

[2]*This method does not provide a measure of the magnitude of deviation of a census statistic from an expected value or a value presumed to be correct, but rather a measure of the variability of census responses attributable to different census operations (for example, interviewers, coders, keyers, etc.).*

[3]*Post-censal matching and reinterview surveys are typically conducted as part of the same post-enumeration survey (PES) operation.*

The reader should study figure 1-2 carefully before continuing with the rest of the manual and return to it after reading the detailed descriptions of each method to obtain a clear picture of how the various methods fit together in an overall census evaluation strategy. Since the various methods measure different components of census error and have different strengths and weaknesses, efficient allocation of evaluation resources must be based upon a thorough understanding of these differences in relation to the information that is sought from the census evaluation effort.

5. *METHODS BASED UPON A SINGLE SOURCE OF DATA*

In the case where the only information for evaluating the census comes from the census itself, methodological options essentially are limited to two choices: (a) the use of selected techniques of demographic analysis, and (b) the use of interpenetration studies implemented as part of the census operation. As might be expected, the "power" of the evaluation methods based upon the results of the census itself is somewhat limited. Fortunately, many developing countries have conducted or will have conducted at least their second census during the 1980 round of censuses. Therefore, some of the more powerful methods based upon the comparison of data from two or more sources can be used to supplement the findings of analysis based solely upon examination of the current census results.

5.1 *Demographic analysis*

When only the results of the census itself are available, a number of techniques of demographic analysis can be used to provide some information on the likely magnitude of error in the data. One of the weaknesses of demographic methods for census evaluation purposes, however, is that they generally do not provide sufficient information to separate errors of coverage from errors in content.

For the purpose of measuring census coverage, only very crude or approximate measures are generally possible. One simple procedure would be to check that each of the smallest geographic units identifiable in the census has a census population total associated with it. While there is no basis for evaluating the accuracy of the total, the lack of a figure for a geographic area assuredly indicates coverage error.

Another approach to examining coverage error on the basis of a single census involves the use of indices which are sensitive to coverage error. For example, one might calculate the average number of persons per household for small areas of the country. Extreme variations from the mean average household size for the country are likely to indicate either coverage error or the existence of special types of residential dwellings (hospitals, barracks, etc.). Comparison of average household sizes for urban and rural areas might also provide some indication of differential coverage since rural households tend to be larger, on average, than urban households in many developing countries.

These are only illustrative of the type of rough coverage checks that could be made using the results of a single census. Other similar types of checks also are possible, but are likely to be as crude as those cited here and this limitation should be kept in mind when using them. As a practical matter, these types of checks are more useful in alerting the users of the data to the likelihood of error than in providing information as to the magnitude or causes of the error(s).

With respect to content error on the other hand, a somewhat more rigorous application of demographic methods to data from a single census is possible involving the analysis of distributions or ratios for particular census characteristics. For example, tabulations of the proportion of ever-married persons by age and sex or of the number of children ever-born to adult females by age could be examined to assess the smoothness of the progression of these cumulative statistics. Irregular progressions or reversals in proportions or means are usually indicative of errors in the marriage, fertility, or age statistics, although it is difficult to determine the relative magnitude of error in each of the characteristics.

Demographic analysts typically make extensive use of sex- and age-ratios in assessing data quality in censuses and surveys. The reason for this is that these ratios "behave" in a rather predictable manner in the absence of catastrophic events such as wars, serious famines or epidemics, and large-scale flows of international migrants. Even for populations affected by these factors, the resulting effects on the age-sex distributions can usually be anticipated and interpreted accordingly.

Analyses of census age distributions can also be undertaken by means of comparisons with "stable" and/or "quasi-stable" age distributions from models. A stable age distribution is defined as the constant age structure that would evolve in a population which has experienced constant fertility and mortality over a fairly long period of time and is "closed" to international migration, while a "quasi-stable" age distribution reflects the effects of declining mortality. While actual populations are unlikely to meet these conditions of stability, quasi-stable population analysis of census age distributions has nevertheless proven to be a quite useful evaluative tool in situations where population age structures have not been affected by wide swings in fertility or international migration.

5.2 *Interpenetration studies*

Interpenetration studies have been used in a number of developed countries to estimate the contribution of one or more census operations to overall census error. What is required to implement this technique is that the assignments of the census personnel whose work is being evaluated (enumerators, coders, key punch operators, etc.) be formed randomly or "interpenetrated." This method is closely related to experimental design methods for testing the effects of varying treatments on groups or samples of subjects. The basic objective of the technique is to assess the extent of variability, or variance, attributable to different census operations. As such, the findings of interpenetration studies are especially useful for the purposes of improving operational control and performance levels in specific operational areas in subsequent data collection activities (for surveys as well as censuses), although the results can also be used in developing measures of accuracy for census statistics from current censuses.

To date, there have been relatively few instances of the implementation of interpenetration studies as part of censuses in developing countries. This may be due in large part to the difficulties involved in controlling these experiments across the various operational phases of a population and housing census or unfamiliarity with the method. With increasing experience and sophistication among statistical organizations in developing countries, however, it is likely that the use of interpenetration studies to identify the operational areas requiring further improvement will become more feasible. Accordingly, a description of the methodology for interpenetration studies is included in this manual (see chapter 3).

6. *METHODS BASED UPON COMPARISONS OF DATA FROM TWO OR MORE SOURCES*

The availability of information from sources independent of the current census opens up a much wider range of choices of census evaluation methods. In addition to there being a larger number of methodological options, the methods based upon comparison of two or more sources of data are considerably more powerful in terms of their ability to assess the relative contributions of different types of errors and, to a somewhat lesser extent, their causes in comparison with methods based upon a single source of data.

As was indicated in figure 1-2, an important distinguishing feature between methods under this heading concerns whether or not they are based upon case-by-case matching of census records with records from the supplementary source. Under the heading of "matching" studies are: (a) post-censal matching surveys, (b) reinterview surveys, (c) record checks, and (d) comparisons with existing household surveys. Methods not based upon record-by-record matching, or "non-matching" studies, include: (a) demographic analysis using previous censuses, (b) comparison with administrative data, (c) comparison with existing household surveys, and (d) interpenetration studies. Each of these methods is described briefly below.

6.1 Matching studies

A common feature of the methods considered under this heading is that they require a matching operation in which individual census records are matched against individual records from another data source to evaluate census coverage and/or the accuracy of the census information for characteristics covered in the census. The methods differ primarily in terms of the type of information which is matched against the census information. Another distinguishing feature is that two of the methods considered (post-censal matching surveys and reinterview surveys) are typically conducted shortly after but in conjunction with the census, while the other two methods (record checks and comparisons with existing household surveys) utilize information which is collected separately from the census, either on a regular basis (population or birth registration systems, for example) or only intermittently (for example, demographic or other socioeconomic household surveys).

6.11 Post-censal matching surveys.--A post-censal matching survey consists of the re-enumeration of an independently-selected probability sample of the target census population and the subsequent determination of "coverage status" in the census on the basis of a record-by-record match. Ideally, an additional sample of census cases also should be drawn and enumerated to check for duplicates and erroneous inclusions in the census enumeration. The results from the sample cases are "weighted-up" to measure coverage for the entire target population. The principles of "dual-system" estimation are typically utilized in making these coverage estimates. (See chapter 2 for a discussion of dual system procedures.)

If properly designed and executed, a post-censal matching survey can provide measures of both gross and net coverage error, as well as information on the components of coverage error (misses, duplications, and erroneous enumerations). For example, coverage error caused by entire housing units or households being missed can be separated from those caused by individuals within enumerated households being missed, as can errors resulting during field enumeration from those occurring during the census processing operations. This information may provide clues as to the reasons for the errors, as well as methods that could prove useful in reducing such errors in future censuses and surveys.

A more familiar name for a post-censal matching survey is a post-enumeration survey, or PES. Strictly speaking, the term post-enumeration survey can be applied to any survey conducted after a census which is used to evaluate the census results. Typically, however, the term is used to describe a sample survey undertaking which is conducted shortly after a census for the primary purpose of evaluating the census. This definition of PES will be used in this manual.

Within the context of this definition, it is important to distinguish between post-censal matching surveys or studies, whose primary objective is to measure census coverage error, and (content) reinterview surveys, whose primary objective is to evaluate the extent of content error in the recorded census characteristics of the population. While the term reinterview survey often is used synonomously with PES, its use in this manual will be restricted to refer to a survey whose objective is to measure content error.

The distinction between post-censal matching surveys and reinterview surveys is an important one for methodological reasons, as well as with regard to the types of errors (that is, coverage versus content) being measured. As is described in detail in chapters 2 and 3, the methodological requirements of the two types of studies are somewhat different. It is possible, however, to combine the two studies into a single post-enumeration survey effort. This has been done quite frequently in both developed and developing countries.

6.12 Reinterview surveys.--As indicated above, a reinterview survey is designed to evaluate the accuracy of the recorded information for selected census items. While discrepancies between the census and reinterview survey information could occur for a variety of reasons (enumerator, coder, or key-punch errors, for example), the objective of reinterview survey evaluation exercises is to measure the broader concept of response error. As might be logically inferred from the description above, since census and reinterview responses to the same census items for the same individuals are required, only reinterview-survey records which are successfully matched to census records (that is, to the census record for the same individual) are considered in assessing the extent of content error.

In describing census reinterview surveys, it is important to distinguish between two types of reinterview survey designs. These designs differ in terms of their measurement objectives. The first type, the reinterview survey with a bias measurement objective, aims at measuring the extent to which the responses recorded in the census differ from the actual or "true" value of the census characteristic for the individual. In order to measure bias, the reinterview survey is designed to obtain more accurate data than were obtainable in the census. It is assumed that more accurate information can be obtained in the reinterview survey through the use of "preferred" data collection procedures, which for any number of reasons were not feasible for use in the census. Preferred procedures could include such things as the use of more highly qualified and better trained interviewers, the use of more extensive probing techniques in soliciting responses to questionnaire items, and (most importantly) field reconciliation of differences between responses given in the census and reinterview situation. To measure "systematic" errors, or bias in the census data, the reinterview survey responses are assumed to reflect the "truth" (or at least are assumed to be more accurate than the census) and thus are used as a "standard" against which the census results are compared.

In the second type of reinterview survey, one with a response variance measurement objective, each person or housing unit is viewed as having a population of responses to a specific question which can be generated by independent repetition of the same survey procedures under the same general conditions. These conditions include such things as the questionnaire used, the method of obtaining responses, the method of recording responses, and the sponsorship of the survey. The census obtains the first of these responses while the reinterview survey obtains the second by applying the same survey procedures under the same general conditions as existed in the census interview. The two responses are assumed to have been selected randomly from the population of responses and are compared to produce estimates of the average trial-to-trial response variability, which is commonly referred to as a simple response variance.

Neither of these two types of studies, in application, can meet its theoretical assumptions. The first type of reinterview survey, one with a bias measurement objective, is unlikely to obtain the "truth" in all cases since the survey is subject to some of the same types of errors as the census (i.e. the respondent may deliberately falsify responses or simply may not know the correct answer). There is also the problem of noninterviews and/or nonresponses to particular items. In the second case, the reinterview survey with a variance measurement objective, the conditions of the original interview may not be duplicated in the reinterview as required conceptually to yield independent responses under the same general survey conditions. For example, a respondent may be conditioned to answer in the reinterview on the basis of the original reply to the census question rather than by attempting to answer the question independently. However, both techniques, if approached carefully, can provide useful information for census evaluation purposes.

6.13 Record checks.--Information collected on a routine basis as part of various types of registration/identification systems often can provide valuable information for use in evaluating censuses of population and housing. A matching procedure similar to that used in evaluating a census on the basis of PES data is used under this approach.

Under this procedure, a sample of records from the registration/record system(s) being used is selected and the relevant persons "traced" forward to the time of the census. The following are examples of the types of record systems which can be used to evaluate census results:

(1) Lists of persons enumerated in a previous census

(2) Registers of births during the intercensal period

(3) Lists of students enrolled in schools, colleges, and universities

(4) Citizen identification cards

(5) Voter registration lists

(6) Registration lists of operators of motor vehicles

(7) Records of national health and social security systems

(8) Records of newly constructed housing units

(9) Registers of households

(10) Immigration register

Since matching is performed on a case-by-case basis, record checks can be used to measure both coverage and content error in the census results. The three major conditions which must be met for record checks to represent a viable option for census coverage evaluation purposes are as follows:

(1) The record system must include clearly defined segments of or the entire target population covered in the census and should be substantially complete; that is, a large proportion of all applicable persons or events should be covered in the system,

(2) The record system must be independent of the census; that is, the probability of an individual being covered by the record system must be unrelated to the probability of that person being counted in the census,

(3) The information provided by the record system should be sufficient to ensure that matching can be performed accurately.

If record checks are to be used for content evaluation purposes, it is also necessary that the record system contain information on at least some of the items covered in the census (for example, age, education, income, etc.) and that the definitions used in the record system be the same or very similar to those used in the census.

In actual practice, these conditions are rarely fully met, particularly the conditions of independence and completeness of coverage of the record system. Despite these limitations, the information resulting from record checks can provide considerable insight into levels and patterns of coverage and content error in population and housing censuses. Among the countries which have used some form of record checks to evaluate censuses are Belgium, Canada, Denmark, Finland, Gibraltar, Guatemala, Honduras, Israel, Italy, the Netherlands, Norway, Sweden, Taiwan, the United States, and Yugoslavia.

6.14 Comparison with existing household surveys.--Theoretically, any scientifically-designed probability sample of housing units, households, or individuals can be used to evaluate a census in roughly the same manner as was described above in connection with post-enumeration surveys. The same basic principles apply to the use of either one-time (ad hoc) or continuing surveys for census evaluation purposes as apply to post-enumeration surveys, the two most important principles being the independence of the survey from the census and the sufficiency of the information collected in the survey to enable the survey records to be matched against the census records to assess coverage in the census. The importance of consistent definitions of characteristics noted above in connection with record checks also applies if existing surveys are to be used to evaluate census content. As was the case with the methods discussed previously, these essential conditions are rarely fully met.

The rationale for the use of household surveys to evaluate censuses in a case-by-case match lies in the fact that, because of their smaller scale, greater operational control can be maintained in comparison with a large-scale undertaking such as a census, with the result that survey data often are thought to be of higher quality than census data. This is more likely to be the case for surveys conducted on a repetitive or continuing basis than for surveys conducted on a one-time or intermittent basis due to more intensively trained and experienced personnel. However, it should be noted that the timing of the survey in relation to the census is a critical factor. With surveys taken a long time prior or subsequent to the census, the difficulties involved in covering the census population (that is, the target population or universe for the census) are likely to be considerable because of the confounding effects of births, deaths, and migration during the intervening period. On the other hand, it is likely to be quite difficult to maintain independence in major survey undertakings conducted shortly before or after a census.

Among the countries which have used existing household surveys for census coverage or content evaluation purposes are the Federal Republic of Germany, Israel, Japan, and the United States.

6.2 Non-matching studies

Non-matching studies, that is those which are not based upon the matching of individual census records with records from another source, range from a quite basic computation of an "expected" census count of population based upon a previous census total and an assumed rate of intercensal growth, to more elaborate procedures based upon separate projections of each of the components of population change (that is, fertility, mortality, and international migration) and the use of administrative data and existing household surveys to derive an expected population estimate. Generally speaking, this group of methods may be described as "demographic methods", although this general concept allows for considerable variability in terms of level of methodological sophistication and quality of results.

In the discussion below, methods are grouped into three categories as follows: (a) demographic analysis using previous censuses, (b) comparisons with administrative records, and (c) comparisons with existing household surveys.

6.21 Demographic analysis using previous censuses.--Methods in this category, as the category heading implies, use information on the size and composition of the population from a previous census along with actual data on or assumptions about the rate of change in the components of population to derive an "expected" population count. This "expected" population is then compared with the population enumerated in the census being evaluated to assess its accuracy.

The most basic of these procedures consists of applying an assumed rate of intercensal population growth to the enumerated population from a previous census to estimate the population expected at the time of the current census. A variant of this procedure would consist of the application of different assumed growth rates for various subpopulations (regions or provinces, for example) to obtain expected population for each subgroup.

Another fairly simple procedure involves the use of the population balancing equation to estimate the expected population at the time of a current census. Under this procedure, the number of births, deaths, in-migrants and out-migrants are added to or subtracted from the population enumerated in a previous census to calculate an expected population size and composition which serves as the basis for evaluating the current census. Of course, this procedure is dependent upon the existence of an essentially complete registration system of births, deaths, and international migration, a target which has not yet been attained in many countries.

In countries with incomplete registration systems, an alternative to the direct use of the population balancing equation is to estimate the number of intercensal births, deaths, and net migrants on the basis of assumptions regarding the level of and trends in fertility, mortality, and migration during the intercensal period. Under this approach, assumed or estimated fertility, mortality, and migration rates are applied to the population enumerated in a previous census to "project" that population forward to the date of the current census. This projected population is then compared to the population enumerated in the current census to assess its accuracy. Assumptions regarding levels and trends in the components of population change can be and typically are developed from the results of demographic surveys conducted during the intercensal period, or even from the previous census using indirect techniques for the estimation of demographic parameters.

6.22 Comparison with administrative records.--In some situations, it is possible to evaluate the accuracy of census results for at least some subpopulations through direct comparison with administrative statistics collected for other purposes. These comparisons yield information on the accuracy of the total count and possibly the composition of the population enumerated with respect to selected important characteristics (for example, age and sex). Examples of the types of administrative record systems which might be considered for use in this manner are records of baptisms, household registers, school enrollment data, and/or social security insurance systems.

6.23 Comparisons with existing household surveys.--The rationale for using existing household surveys to evaluate a census on a non-matching basis (that is, where no attempt is made to match individual census and survey records) lies in the fact that sample surveys are often less affected by nonsampling error than censuses. As noted previously, this is due in large part to the greater operational control that can be maintained in a sample survey situation in comparison with a census. As a result, survey data on population characteristics also measured by a census often are considered to be more accurate than the census results.

The use of existing household surveys on a non-matching basis for census evaluation purposes is limited by two primary factors. First, unlike a complete enumeration, the survey estimates are affected by sampling error which must be taken into consideration when comparing the census and survey data. Second, the timing of the household survey being considered in relation to the census is a critical factor. If the reference date for the survey is too far removed from the census date, comparisons between the two sets of estimates will include differences arising simply because the population size and composition with respect to important characteristics had changed during the intervening period. In such a case, it is difficult to separate errors from actual changes in important characteristics of the population being studied. These factors significantly limit the use of household surveys for evaluating censuses on a non-matching basis, although some useful information can be obtained if the analyses are approached carefully.

7. *RELATIVE STRENGTHS AND WEAKNESSES*

Each of the methods described in the previous sections of this chapter has strengths and weaknesses associated with it. In considering the relative merits of the different approaches, the following criteria should be taken into account:

(1) The types of errors which are measureable by each method;

(2) The number and types of assumptions which must be made in order to apply the method;

(3) The degree of difficulty typically encountered in satisfying the necessary assumptions and the consequences of failure to do so;

(4) The level of technical and financial resources required to implement the complex operations.

Since methods classified under each of the general approaches outlined in figure 1-2 tend to share the same strengths and weaknesses, the discussion below is organized in terms of comparisons among these general approaches. Strengths and weaknesses specific to each method are also indicated, however.

7.1 *Methods based upon single versus multiple sources of data*

Generally, methods based upon a single source of data (that is, the census itself) provide less insight into the magnitude and types of error present in the data than do those based upon comparisons of two or more sources of data. For example, age and sex distribution analyses, either by means of analyses of age-sex ratios or stable/quasi-stable population analysis, provide a general impression of the quality of the census results, but provide little insight on the relative contributions of coverage and content error or on the issue of bias versus variance. These types of analyses tend to be most useful when the results can be compared with results obtained by other methods (for example, estimates of net coverage error by age derived from a PES or other type of matching study).

An advantage of this type of evaluation study, of course, lies in the fact that it does not require additional data to be collected for evaluation purposes, nor are sophisticated designs and time-consuming matching operations required. Since many national statistical offices in developing countries have qualified demographic statisticians on their staffs, there is often no need for significant increases in staffing levels or for technical assistance in carrying out this type of evaluation. In view of these factors, it is generally desirable to carry out this type of evaluation exercise even when other methods also are used in order to provide a more comprehensive picture of the quality of census data.

Interpenetrating surveys, if properly designed and executed, can provide considerable insight into the relative contribution of component errors at different operational stages to total census error. Their primary use lies in the identification of the operational areas in which improvements are needed for future censuses. These surveys also provide measures of variance associated with various operational stages for the current census which can be used as input in calculating measures of accuracy for selected census items (see chapter 3 for further details). They do not, however, provide evidence on the relative magnitude of coverage versus content nor net versus gross error, and thus may be most useful in countries with previous experience in census evaluation and/or in conjunction with other evaluation methods.

The United States has used interpenetration studies with much success since about 1950. More than any other content evaluation study, the interpenetration study has shaped the design of the census. The method has a number of drawbacks, however.

Interpenetration study design and analyses require technical personnel with advanced statistical skills, a serious disadvantage when such resources are scarce. From an operational perspective, the field procedures can be made relatively simple and can be carried out by the regular field staff with some additional training. However, more record keeping is required to record the unique enumerator identifier associated with every census questionnaire in the study sample. Moreover, additional supervision is required to ensure that procedures are being followed, i.e., the relevant information is captured and the study enumerators are maintaining their interpenetrated assignments. Periodic site visits by the project leaders are mandatory.

From a cost perspective, the major disadvantage is the increased enumerator travel costs, since study enumerators cover twice the area covered by the conventional enumerators. Of course, the additional training, record keeping, and supervision also add to the cost increases.

Despite these disadvantages, interpenetration studies provide useful information which is available by no other means. For small areas, especially, the correlated component of enumerator variance is believed by many to be the major component of nonsampling variance. The degree of control required to execute these studies, and the level of statistical sophistication required to analyze the resulting data may discourage many countries from tapping its true potential.

7.2 *Matching versus non-matching studies*

The primary advantage of matching over non-matching studies lies in their ability to provide separate estimates of various types of error; that is, coverage versus content and net versus gross. Non-matching studies, because they consider the census results at the aggregate level rather than at the level of individual housing units, households, or persons, provide only estimates of net census error. Further, it is often difficult to discern the relative importance of coverage and content error from non-matching studies, and in many cases conclusions on this matter must be based upon the judgement of the analyst and/or upon indirect rather than direct evidence.

Where at least two censuses and reasonably accurate information on levels of fertility, mortality, and migration are available, demographic analyses can provide defensible and consistent estimates of census coverage (at least at the national level) and substantial evidence on the overall quality of census age data. However, since estimates of census error are derived as "residual" differences between the actual and expected census counts, it is important to have fairly accurate information on levels of fertility, mortality, and migration. The accuracy of estimates of census error derived from demographic analyses of successive censuses depends entirely upon the accuracy of the information from the previous census and on the components of population change. Where this information is of uncertain quality, it is often difficult to determine what portion of the estimated census error to ascribe to errors made during the census being evaluated as opposed to errors attributable to the data used in the calculation of the expected population.

Aggregate-level comparisons of census results with administrative data and existing household surveys, on the other hand, are affected by coverage and completeness of the data being compared to the census. In the case of administrative data, an assessment must be made as to the relative completeness of these data as well as to the net effect of differential biases which might exist in the two sets of data being compared. With regard to household surveys, both the effects of sampling error in the survey estimates and the extent of possible correlation bias between errors in the census and the survey need to be taken into account. The latter problem is often inevitable when a previous census is used as a sampling frame for the survey and coverage errors in the two censuses are correlated.

While being able to provide separate estimates of coverage and content or net and gross error, different requirements apply in the implementation of matching studies. The most readily apparent of these is the significantly higher levels of technical, managerial, and financial resources. The additional resources required to design and implement a matching study and analyze the results are common to each of these methods. In the case of post-enumeration surveys, an additional round of data collection also has to be undertaken prior to the matching operation.

Aside from these greater operational demands, the major difficulty arising in the implementation of matching studies involves the extent to which the theoretical assumptions of the underlying methodology can be satisfied in actual practice. Typically, biases of several types are present (and to a large extent unavoidable). Response correlation bias results from the fact that the two data systems being compared are not in fact independent; that is, the probability of a subject being covered in the census is correlated with the probability of being covered in the second data system. Matching bias results because the rules and information used in conducting matching operations are imperfect and errors are made in determining "match-status." Out-of-scope bias arises when persons who should not have been are included in either system of data being compared.

In general, correlation bias tends to be relatively low in the case of record checks since the causes of omissions and erroneous inclusions are substantially different than those encountered in censuses and surveys. Matching bias, on the other hand, can be more readily controlled in a PES than in other matching methods because the data collection procedures can be designed specifically to facilitate matching.

Considerable effort must be made in the conduct of matching studies to minimize these biases to the extent feasible. To a large degree, the quality of results of matching studies are determined by the extent to which these biases can be minimized, although bias can never be eliminated fully in actual practice.

Finally, it should be noted that the range of census characteristics which may be evaluated with a matching study is considerably larger than is generally possible in demographic analyses, which are often demoted to the assessment of the accuracy of census age and sex data.

7.3 *Basis for choice of methodology*

Due to the unavailability of sufficiently reliable record and/or registration systems, the primary methodological options for census evaluation purposes in many countries are demographic analyses and the implementation of a post-enumeration survey.

Since demographic analyses are typically undertaken in connection with censuses irrespective of whether a PES is conducted, the critical decision often comes down to whether or not to conduct a PES. This decision must be based upon judgement as to whether the cost of a PES can be justified on the basis of the additional information to be gained.

An important factor in this decision concerns the quality of demographic data available. Since demographic methods for estimating census errors, particularly coverage errors, are not very reliable unless reasonably accurate information is available on fertility, mortality, and migration levels and trends, the absence of reliable information should dictate caution in adopting demographic studies as the sole basis for evaluating census coverage. This caution is especially relevant in countries where the level of international migration is substantial, since migration data are almost always of lower quality than are data on fertility and mortality.

In the situations of many developing countries, the PES approach may be the only reliable means of evaluating census errors, particularly coverage error. The difficulties involved in controlling biases in a PES should, however, be borne in mind. In particular, several developing countries have experienced difficulty in controlling response correlation bias, resulting in estimates of census error which were thought to represent a "lower bound" of the actual degree of error in the census.

Nevertheless, an increasing number of developing countries have undertaken PES evaluation studies in recent years. An indication of the extent to which PES has been used to evaluate census results in developing countries is provided in figure 1-3.

Figure 1-3. DEVELOPING COUNTRIES IN WHICH POST-ENUMERATION SURVEYS WERE CONDUCTED IN CONJUNCTION WITH THE LATEST CENSUS

Region/ Country	Year of Census	Type of Error Measured	
		Coverage	Content
AFRICA			
Algeria........	1977	x	x
Botswana.......	1981	x	-
Burkina-Faso...	1975	x	x
Burundi........	1979	-	x
Cameroon.......	1976	x	-
Camoros........	1980	x	-
Guinea-Bissau..	1979	Planned	Planned
Ivory Coast....	1975	x	-
Madagascar.....	1975	-	x
Malawi.........	1977	x	-
Morocco........	1982	x	-
Seychelles.....	1977	-	x
ASIA			
Bangladesh.....	1981	x	x
Burma..........	1981	x	x
China (P.R.C.).	1982	x	x
Cyprus.........	1982	x	-
India..........	1981	x	x
Indonesia......	1980	x	-
Korea..........	1980	x	x
Malaysia.......	1980	x	x
Pakistan.......	1981	x	x
Philippines....	1980	x *	-
Sri Lanka......	1981	x	x
Thailand.......	1980	x	x
Yemen (Sanaa)..	1975	x	-
LATIN AMERICA			
Argentina......	1980	x	x
Bolivia........	1976	x	x
Brazil.........	1980	x	x
Colombia.......	1973	x	-
Cuba...........	1981	x	x
Mexico.........	1980	x	x
Uruguay........	1975	x	x
Venezuela......	1981	x	-

**Metro Manila only*

All countries which conducted some form of PES in conjunction with their most recent census (based upon available information) are listed in figure 1-3, along with the year of the census and type of error studied (coverage/content). As indicated, the number of countries undertaking PES studies in connection with their last census is substantial, although it should be noted that the PES designs implemented in these countries are subject to considerable variability. Several other countries (Niger and Senegal, for example) conducted "control" surveys in connection with their most recent censuses (1977 and 1976, respectively), but these were implemented apparently more for quality control than error measurement purposes and as such were not included in figure 1-3. In addition, several countries (Colombia, Egypt, Senegal, Somalia, and Yemen) were preparing for the implementation of a PES in connection with their 1985 or 1986 censuses at the time of preparation of this manual.

Profiles of PES studies with coverage and content error measurement objectives for selected countries are provided in chapters 2 and 3.

8. REFERENCES

Cochran, W.G. 1953. *Sampling Techniques.* New York: J. Wiley and Sons.

Hansen, M.H.; Hurwitz, W.N.; and Madow, W.G. 1953. *Sample Survey Methods and Theory.* New York: J. Wiley and Sons.

Kish, L. 1965. *Survey Sampling.* New York: J. Wiley and Sons.

Marks, E. 1978. "The Role of Dual System Estimation in Census Evaluation." In *Development in Dual System Estimation of Population Size and Growth*, edited by K. Krotki. Alberta: University of Alberta Press.

Chapter 2. STATISTICAL EVALUATION OF COVERAGE ERROR

1. INTRODUCTION

No matter how carefully a statistical office conducts a census program, errors in the coverage of the census are inevitable. Two types of coverage error will be discussed in this chapter: undercoverage and erroneous enumerations. Undercoverage in a census results from erroneously omitting a person, housing unit, or household. Severe undercoverage can lead users of the census data to question the validity or even utility of the results. If the rate of undercoverage is different among major population subgroups, the census can present distorted statistics for the group the data represents.

Normally less frequent, but nevertheless of concern, are erroneous enumerations in the census. Erroneous enumerations are defined as persons, housing units, or households that were enumerated when they should not have been, or enumerated incorrectly. This includes duplicate or multiple enumerations, such as enumerations that should not have occurred because the persons or housing units do not exist (called "curbstone" cases) or enumerations that are misassigned according to geographic or demographic subgroup. Duplicates or curbstone cases will lead to incorrect census totals at all levels in the statistics tabulated from the census, while enumerations that are incorrectly assigned according to geographic or demographic subgroups will be correctly represented at the highest levels of statistical aggregation, but may introduce biases into the estimates of census coverage error.

This chapter presents alternative methodologies for the estimation of coverage error in censuses of population and housing. More specifically, attention is directed to studies based upon direct matching with census records to determine who was and was not enumerated in a particular census. Nonmatching studies are described in chapters 4 and 5.

In view of its wider applicability to developing country settings, the major emphasis of the material presented in this chapter is on the methodology of a post enumeration survey (PES). Following this, alternative methodologies based upon longitudinal tracing studies and the use of network (multiplicity) sampling are described briefly. The chapter concludes with a series of short profiles of census coverage studies which have been undertaken in both developed (the United States and Canada) and developing countries (Korea, Paraguay, India, and Bangladesh). The latter material is included in order to provide some indication as to alternative study designs which have been attempted under varying conditions, as well as some of the problems which have been encountered under these designs.

2. POST ENUMERATION CENSUS MATCH STUDY

A post enumeration survey (PES) census match study generally serves four purposes:

(1) The PES can indicate to data users where specific coverage problems occur in the census data and quantify these errors;

(2) The PES can provide guidance to census planners in designing future censuses, i.e., efforts can be made to improve coverage in difficult-to-enumerate subgroups of the population;

(3) Aside from identifying difficult-to-enumerate subgroups of the population, the PES can identify problem or erroneous procedures used in the census. Since the PES is a case-by-case match study,

situations which lead to units being procedurally missed or erroneously enumerated will become apparent;

(4) If the data problems in the census are sufficiently severe to warrant an adjustment of the census, the PES can provide detailed information to be used in adjustment models. Adjustment technologies will be described in chapter 6. A later portion of this chapter will discuss the sample design of the PES, and how the sample can be designed to facilitate an adjustment.

The post enumeration survey actually consists of two separate coverage studies. The key elements are:

(1) A survey conducted using a sample drawn from a sampling frame independent of the census being evaluated. Persons and households from this survey are matched to the census to estimate the number of persons missed in the census.

(2) A survey conducted using a sample drawn from persons enumerated in the census. This sample is reenumerated to determine if the sample person or unit was erroneously enumerated. Estimates of erroneous enumerations include counts of persons or units that are duplicates in the census, persons or units that should not have been enumerated in the census, and persons or units that are incorrectly located according to geography (misplaced geographically).

The next section of this chapter describes a statistical model used to combine the results of the two surveys and to generate estimates of the net undercount in the census. Succeeding sections describe how misses and erroneous enumerations are defined and measured.

2.1 *Dual system estimation*

To develop estimates of coverage combining the results of the two surveys, one must use a statistical model. The model indicates how errors in the census occur as a stochastic process and how an appropriate estimator of census undercoverage may be derived. This section presents a simple derivation of the dual system estimator, followed by a modification of the estimator to allow for erroneous enumerations in the census and measurement problems in the PES.

2.11 Modeling census misses.--The model for coverage evaluation was originally developed for use in biometric studies to estimate the size of closed populations. A closed population in the biometric sense is one in which the composition of the population remains relatively unchanged over the time the study is being conducted: there are no births or deaths, and there is no immigration or emigration. The earliest uses of the technique, known in biometrics as capture-recapture estimation, was for the purpose of estimating the sizes of wildlife populations. An example is given to indicate how the model was originally derived.

Consider a lake with no inlets or outlets. It is desired to estimate how many fish are in the lake. Assume the fish are uniformly and randomly distributed throughout the lake. We take a net and "capture" as many fish as possible in one catch. The fish are counted (and the total denoted N_1), tagged, and released back into the lake. The tags do not injure or affect the fish in any way. After sufficient time has passed for the fish to redistribute themselves randomly throughout the lake (but not so long a period that one might encounter births or deaths of any fish), another net is cast into the lake and a second catch is taken. These fish also are counted and denoted N_2. A separate tally is also made of fish tagged during the first catch who were "recaptured" in the second catch, denoted M. We now have three totals, and the captures can be distributed as in figure 2-1 below.

Figure 2-1. DISTRIBUTION OF CAPTURES IN A CAPTURE - RECAPTURE STUDY

First Capture	Second Capture Total	Caught	Not Caught
Total.......	N_T	N_2	
Caught........	N_1	M	
Not Caught....			

where

N_1 = the number of fish caught in the first attempt

N_2 = the number of fish caught in the second attempt

M = the number of fish caught in both attempts

N_T = the total number of fish in the lake

The objective of the exercise is to estimate N_T, the number of fish in the lake; but the number of fish not caught in either attempt remains unknown to this point. To model the capture process, the capture of a fish is conceptualized as a stochastic event, or more specifically, a Bernoulli event. This implies that the counts of fish caught, M, N_1, and N_2, are random variables (the sums of Bernoulli outcomes). If the probability that a fish will be caught in the first attempt is "a" for each fish, and "b" in the second attempt, and the two captures are independent of each other, the expected values of the random variables M, N_1, and N_2 may be evaluated as follows:

(2.1) $E(M)=abN_T$ (by independence of events)

(2.2) $E(N_1)=aN_T$

(2.3) $E(N_2)=bN_T$

An estimator may be obtained by substituting the observed data values for the expected values of M, N_1, and N_2, and substituting equations (2.2) and (2.3) into equation (2.1). Thus,

$$(2.4)\quad M = abN_T = \frac{aN_T\,bN_T}{N_T} = \frac{N_1 N_2}{N_T}$$

Rearranging terms in (2.4) and we have:

$$(2.5)\quad N_T = \frac{N_1 N_2}{M}$$

2.12 <u>The model applied to censuses under perfect conditions</u>.--In a census evaluation application, the theory is exactly the same as the capture/recapture study described in section 2.11, except that human populations require slightly different assumptions. Suppose that the first capture is the census and the second capture is the PES. Let N_c (replacing N_1) be the census count, and N_p (replacing N_2) be the weighted sample total from the PES. N_p represents the total number of people (or units) to be found if the entire sampling frame used for the PES sample were contacted. It is critical, of course, that the sampling frame is independent of the census. For estimation purposes, it is best to have as complete a frame as possible so as to minimize the variance of the estimate and the number of persons (units) not found in either data source. The random variable M represents the number of people (units) "captured" in the PES who were enumerated in (are matched against) the census. The estimate is made in exactly the same way as in equation (2.5).

Note that there are a number of assumptions that have to be satisfied in applying the model to human populations. The model assumes that matching of persons and units is done perfectly between the two sources, that there are no multiple observations of individuals in either source, and that the two sources are independent. This latter assumption must be built into the design of the study by keeping the PES as independent as possible from the census. Therefore, the frame for the PES sample must be independent from the census. Perhaps the best choice for a sampling frame would be a current updated listing of the previous census or an area sample based on a recent listing of areas. Other requirements for keeping the two sources independent are using interviewers for the PES who are different from those used for the census and processing the PES separately from the processing of the census to avoid cross-contamination. The PES sample used to estimate census misses will be referred to as a "*P*-sample" in the following discussion in order to differentiate it from the evaluation sample drawn from the census, which shall be referred to as the "*E*-sample."

2.13 The model applied to censuses under less than perfect conditions.--There are several types of errors that can affect the estimation procedure described above:

(1) The census may contain duplicate or multiple enumerations.

(2) Housing units listed in the census may be correctly enumerated but allocated to the wrong geographic area.

(3) Members of a housing unit may be enumerated in the wrong location or may not be in scope for the census (i.e., should not have been counted at all).

(4) Members of a housing unit may be incompletely enumerated so that there is insufficient identifying information for an individual.

Any of these factors can introduce bias into the estimation of the total population size. To measure or counter the affect of each of these factors, a second sample is used to provide correcting factors to the estimation process. This second, or *E*-sample, is drawn directly from the census for the same area and using the same stratification as the sample described in the previous section.

A followup enumeration is conducted at each of the households in the *E*-sample and several pieces of information are collected. When the sample housing unit is initially contacted, the geographic location is checked and compared to the geographic location recorded for that housing unit in the census. If the locations differ, the case is reviewed (usually by a third party) and a determination is made as to whether the original geographic coding in the census was correct or incorrect. The reason for checking geographic coding is that housing units incorrectly located in the census is unlikely to be found when the PES match is conducted, unless unlimited resources are available to search. Thus, although the housing unit is correctly enumerated in the census, none of the occupants can be found or matched. As a result the census total, N_c, will be correct, but the match total, M, will be too low. The method to correct the match total is presented in section 2.14.

In the enumeration undertaken for this sample, the interviewer ascertains for each occupant of the sample housing unit whether the occupant listed in the census actually exists (there are sometimes cases which are fabricated in the census by overzealous census enumerators) and whether that individual was correctly enumerated at that address in the census. A count of the number of persons who should not have been enumerated at that address is obtained for each household. In the case of census fabrications, these "non-persons" should be subtracted from the census totals since they don't exist. In the case of incorrect enumeration, if the PES interview was conducted correctly following the census enumeration procedures, the PES interview will either not list the individual (out-of-scopes) or have them listed at another address, and in neither case will there be a match. In this case the census totals, N_c, are incorrect and there should be no matches.

There are two further operations to be conducted in the PES processing office. For the area of search specified for the match portion of the PES operation, the sample cases drawn from the census should be compared to all other census returns in the same area. This comparison should be done for each person listed on the sample unit return to determine whether there are any duplicate enumerations in the same area. Duplication in the census will cause the census total, N_c, to be overstated, though the number of matches made from the PES sample, M, will be correct.

Finally, there will be cases in the census which have insufficient information for matching to the PES sample. Although tallied into the census (probably as a result of imputing much of the missing information) so that the census total, N_c, is correct, the match cannot be completed and the match total, M, will be understated.

There are four totals that can be obtained from this second sample and from the computer files of responses. They are designated as:

- G the number of persons incorrectly located geographically in the census
- E the number of persons incorrectly enumerated in the census (fabricated or not in scope)
- D the number of duplicate enumerations in the census
- I the number of persons who are enumerated in the census but have insufficient information for matching

2.14 <u>The estimate of the total population size and the net undercount rate</u>.--Equation (2.5) provides the estimator of the total population size when the data used in the capture-recapture study are without error. When errors are present, the census totals have to be adjusted to remove duplicate enumerations, fabrications, and cases which have no chance of being matched. The final estimator of the total population size is

$$(2.6) \qquad N_T = \frac{N_p \cdot (N_c - G - E - D - I)}{M}$$

The net undercount rate, R_n, can be calculated as the total census count relative to the estimate of the total population size, or

$$(2.7) \qquad R_N = \frac{N_c}{N_T} = \frac{M/N_p}{(N_c - G - E - D - I)/N_c}$$

2.2 *Alternative procedures for misses and erroneous enumerations*

The method for conducting a matching operation and forming an estimate depends on the survey procedures and definitions chosen for the PES. There are three basic procedures that can be used in a PES to evaluate coverage in a census. The procedures differ in the treatment of "movers"; that is persons whose location at the time of the PES differ from their location at the time of the census.

2.21 <u>Procedure A</u>.--This procedure reconstructs the households as they existed at the time of the census. A respondent is asked to identify all persons who were living or staying in the sample household on census day. These persons are then matched against names on the census questionnaire for the sample address. From this information, estimates of the number and percent matched for non-movers and out-movers (persons who have moved away since the census date) can be made.

2.22 Procedure B.--This procedure identifies all current residents living or staying in the sample household at the time of the PES. The respondent is asked to provide the address(es) where these persons were living or staying on census day. These persons are then matched against names on corresponding census questionnaire(s). Estimates of the number and percent matched for non-movers and in-movers (persons who have moved into the sample address after the census date) can be made.

2.23 Procedure C.--This procedure identifies all current residents living or staying at the sample address at the time of the PES plus all other persons who lived at the sample address on census day. However, only the census day residents (non-movers and out-movers) are matched with the census questionnaire(s). Estimates of the number of non-movers, out-movers, and in-movers, and of the percent matched for non-movers and out-movers, can then be made. Estimates of non-movers and movers come from Procedure B and match rate estimates from Procedure A. Thus, Procedure C is a combination of Procedure A and B.

The difference between Procedure A and Procedure B relates primarily to movers. For non-movers, theoretically one should get the same listing whether one lists the people who were living in the housing unit (HU) on Census Data, as in Procedure A, or the people who are living in the housing unit at the time of the PES, as in Procedure B.

In Procedure A, persons in the sample segments are asked about out-movers. Present residents and neighbors often do not know the former occupants; thus, Procedure A usually leads to a large undercount of movers. On the other hand, since out-movers should have been enumerated in the EA's in which the PES sample is taken, the matching of out-movers is relatively easy and inexpensive.

Procedure B enumerates in-movers at their new addresses in the PES sample segments. Thus in-movers can supply information about themselves. Consequently, Procedure B should give a better estimate of movers, even when in-movers give poor addresses for their former residences. However, finding the census records for the former addresses is a very difficult task. Often there is insufficient information for locating the census EA in which the mover should have been enumerated. Thus, even though Procedure B may give a better count of movers than Procedure A, the problems and costs of matching the PES and census records are considerably greater than in Procedure A.

Procedure C combines features of Procedure A and B with the objective of reducing the matching problems while taking advantage of a better count of movers. In Procedure C, in-movers are used to get the count of movers, both total and subgroups by sex, age, relationship, etc. The out-movers, as in Procedure A, are matched for the purpose of estimating the proportion enumerated and the proportion missed in the census. The proportions enumerated and missed are applied to the count of in-movers to get the counts of enumerated and missed movers; these counts are added to the corresponding estimates for non-movers.

2.24. Advantages and disadvantages.--Since the three procedures differ with respect to the treatment of movers, their relative advantages and disadvantages are also in this area.

Procedure A. The weakness of Procedure A is in the fact that movers are no longer at the sample address. In a majority of cases, no member of the mover's family is living at the sample address and the mover must be reported by a proxy respondent. There is often considerable difficulty in finding a satisfactory respondent. Trying to get the out-mover's new address is difficult and expensive. Even if an acceptable respondent can be located at the old address, enumerators report considerable difficulty in obtaining the required information.

The enumeration problems of Procedure A are considerably amplified for a de facto census. The enumerator must ask about people "who were staying here at the time of the census" including, at least in theory, one-time visitors and persons who may have slept in the street near the house without the occupant's knowledge.

Procedure B. The field procedures for Procedure B are simpler than those of Procedure A and easier for the enumerators since (a) in most cases, information can be obtained from the person or from his/her immediate family and (b) it is easier to make the respondent understand the questions. Procedure B tends to get more complete and more accurate answers.

Against these advantages in the field procedures for Procedure B, there are serious disadvantages in matching. It must be remembered that a "match" is proof that the person was enumerated in the census, but a "non-match" is proof only that the person was not enumerated in the areas searched. With Procedure B, one is never sure whether failure to find a match in the census indicates a missed person or an incorrect or ambigous address.

In matching the in-movers in Procedure B, it is necessary to find the old address in the census questionnaires. Particularly with rural addresses, the indicated address may be in one of four or five (or more) EA's which may contain several hundred households. In many cases, Procedure B may require examining a number of census questionnaires for a possible match. While this happens in Procedure A also, one can search the questionnaires for all the households in a Procedure A segment at one time.

Procedure C. Procedure C combines the relatively simple and inexpensive matching operation of Procedure A with the more straightforward and complete field enumeration of Procedure B.

Procedure C will be slightly more expensive than Procedure A since it requires enumeration of in-movers in addition to the out-movers and non-movers covered by Procedure A. However, Procedure C will have considerable advantages over Procedure B in the costs and difficulties of matching, while the incremental field costs of adding an enumeration of out-movers to the PES field procedures should be small.

Compared with the two other procedures, C would seem to give results with lower correlation bias than A, lower matching bias than B, and probably somewhat lower overall bias than either A or B. It is expected also that C would give results with a variance somewhat higher than B and slightly lower than A.

2.3 *Alternative definitions for misses and erroneous enumerations*

In addition to the procedures, it will be necessary to define the area of search (geographically) where an individual or housing unit is expected to be found based on information obtained in the PES interview. Two sets of definitions can be used for this purpose. Definition I gives a very broad area where a person or unit might be found, whereas Definition II limits the searching operation to one area where the respondent should be found. Definition I will naturally yield higher estimates of census matches, but this will be offset to some extent by higher estimates of geocoding errors and erroneous enumerations from the *E*-sample. Definition I will seem to yield a lower undercount rate than Definition II, although with use of the *E*-sample, both estimates have the same expected value. Definition II will be less expensive to implement because of reduced searching for possible matches, but the estimates will have a higher variance than estimates under Definition I. Definitions I and II are summarized in figure 2-2.

For Definition I, the search for misses is done throughout those EA's that correspond to the reported alternative locations where the person might have been enumerated. Under Definition II, the search is limited to the EA where the PES reported the person should have been enumerated. Thus, for Definition II, the person is erroneously enumerated only if the census day residence reported in the PES is in a different EA than was reported in the census.

2.31 Advantages and disadvantages of Definition I.--All locations where a person might have been enumerated in the census must be obtained by the PES and a search made of the census lists in order to classify the person as correctly enumerated, missed, or erroneously enumerated. Obtaining all addresses where the person might have been enumerated

Figure 2-2. COMPARISON OF DEFINITIONS I AND II

Classification	Definition I	Definition II
Correctly enumerated..........	Enumerated only once (even though the person was enumerated at the wrong location) in a location where the PES reported the person might have been enumerated.	Enumerated in the location where the PES reported the person should have been enumerated.
Missed........................	Should have been enumerated in the census but was not in any location (correct or incorrect) where the PES reported the person might have been enumerated.	Not enumerated at the location the PES reported the person should have been enumerated. The person may have been enumerated somewhere else in the census while the PES reported the true location, or the PES may have been incorrect in reporting the persons location on census day.
Erroneously enumerated........	Enumerated more than once* or should not have been included in the census; for example, a person who was "invented" by the census enumerator, a person born after census day or a person who died before census day.	Reported by the census erroneous enumeration sample as not living at the census reported location.

**If found in the census records in more than one location, the person is counted as correctly enumerated in one location and erroneously enumerated in the others.*

is very difficult for Procedure A out-movers; proxy respondents will not be able to always provide this information. An interview to obtain information from Procedure A out-movers can be expensive and often ineffective. Thus, Definition I is usually restricted for use with Procedure B. However, even with Procedure B, Definition I has problems. It is extremely difficult to obtain alternative addresses. Moreover, the matching of alternative addresses is expensive and often quite difficult. Persons who identify with many "locations" may be reluctant to divulge that information.

The basic relative strength of Definition I is one of relatively small variance and bias, which is an important consideration in making coverage estimates for small geographic areas and for specific subgroups of the population; this is due to the "residence at time of census" concept that is elaborated on in the discussion of Definition II in section 2.32. Also, a separate sample of census enumerations for a searching operation to estimate census duplicates is not needed, in contrast to Definition II; resulting in cost savings. The advantage of low bias, of course, assumes that relatively little response error is associated with obtaining alternative locations where the person could have been enumerated.

2.32 Advantages and disadvantages of Definition II.--The weakness of Definition II concerns the "residence at time of census" concept. Persons enumerated in one place in the census and reported by the PES as living elsewhere are counted as both "missed" and as "erroneously enumerated" if the erroneous enumeration sample is overlapped with the *P*-sample. This can happen frequently and will increase the variance of the estimate. Bias can be introduced by the inconsistency of the response to the "place of residence at time of census" question. A tendency could exist for Procedure B in-movers to report their current sample address as their census day address. Very often this is the "easiest" response to the question. However, if this person was included in the sample of census enumerations used to measure erroneous enumerations (at their former address), a proxy respondent at the census day location could possibly report the person as being enumerated at that address in the census. This, of course, assumes that there would not be a followup of out-movers for the sample to measure erroneous enumerations due to the expense and difficulty of obtaining forwarding addresses.

2.4 *Survey and sample alternatives*

It should now be obvious that conducting a PES entails choices among many alternatives. In the design of the study, one can choose among Procedures A, B, or C to determine how the interview will be structured. A second choice which must be made is how the match will be conducted; that is, using Definition I or Definition II. The choice of procedure and definition will have an impact on the total survey design by directing the development of the questions and dictating how the match to the census will be conducted. To make a choice that will prove most viable for the census being evaluated, the procedures and definitions to be used should be pretested in the field well before the census is conducted. Pretesting gives an indication of the types of problems to be expected in the actual PES and allows sufficient opportunity to modify the study design. If more than one PES procedure is tested, these procedures can be compared and the more appropriate design chosen.

After the 1950 Census of Population in the United States, a post enumeration survey was conducted to measure coverage error. An intensive reenumeration of a sample of areas was conducted, attempting to find all cases missed by the original enumeration. Cases within the sample areas were matched to the census and differences which appeared between the PES and the census were dependently reconciled. The most highly qualified enumerators were used for the PES and were extensively trained. A subsequent PES (for the 1960 census) attempted to improve upon the 1950 experience by use of more probing questions and by a more thorough canvassing of the sample areas. In general, the results of these two post enumeration surveys were disappointing. The amount of underenumeration detected was significantly less than estimates produced using demographic methods of census evaluation.

Successful experimentation with dual system estimation techniques for evaluation of coverage of vital registration in the United States during the 1960's suggested that a new approach may be warranted. This new approach emphasized independence of the PES and the census in addition to better quality and improved enumeration techniques. The older PES approach recognized that if a census is poorly done, a PES done more carefully can pay dividends in revealing census error. However, there is a point where increasing the expenditure for quality will meet with decreasing returns in providing better coverage error estimates. A new approach (developed during the 1960's and 1970's) also advocates having quality for the PES. Probing is encouraged, as well as thorough training for enumerators. However, the major emphasis is on keeping the PES operations independent from the census. Some suggestions for maintaining this independence are:

(1) Wait until all census questionnaires have been returned from the field before conducting the PES field enumeration.

(2) Attempt to ensure that PES areas are canvassed by enumerators other than the census enumerators. Do not let PES enumerators know what was enumerated in the census.

(3) Conduct PES processing separate from and independent of the census processing.

In terms of sample design for the PES, an initial decision must be made whether to develop a sample specifically for the purpose of census evaluation or to conduct the PES as a supplement to a survey scheduled to be undertaken shortly after the census in which demographic and socioeconomic data will be gathered. Examples of these are labor force, demographic, or health surveys. The relative advantages and disadvantages of these two alternatives and some general design considerations are indicated below.

2.41 <u>Specially designed sample for PES *P*-sample</u>.--If measurement of coverage is an important concern for the evaluation of the census, a sample specifically designed for the PES will offer several advantages over use of a sample that was designed for multiple or other uses. Standard sampling theory would indicate that a sample designed specifically for coverage would allow optimum allocation of the sample so as to achieve a fixed upper bound for the variance of the estimate for a minimum sample size. The variance of the estimate can be reduced by stratifying the sample using estimates of the undercount from previous censuses or using variables suspected of being correlated with the undercount.

Two special considerations enter into the design of a sample for measurement of undercoverage that usually would not be part of the sample design decision process. The first is that the sample should be designed to facilitate the matching of the household listings to the census. The second is that, if an adjustment is to be made, there must be a basis for allocating the measured undercount to areas not in sample, or to small geographic units (problems of adjustment are discussed in chapter 6). A possible resolution to both problems is to design the last stage of the PES sample to be a block sample.

In the U.S. census, most portions of the country are divided into blocked areas. A block is defined as a relatively small geographic area uniquely bounded by streets, railroad tracks, rivers or streams, or other natural boundaries. The advantage to using blocks is that the households contained in a block form a closed set. If the entire block is sampled, there is no question about which units are listed for a PES or fall into the sample. Since the census is organized by blocks, it is relatively easy to match a block listing of housing units from the PES to a block listing from the census. Units in both listings can be defined in terms of their location relative to one another, especially if both listings have accompanying spot maps. In addition, by sampling blocks as they are defined in the census, the geography for each case is settled, thus minimizing the number of geographic placement errors in the PES and the cost of assigning geographic codes.

If the census being evaluated does not divide or collect census materials by block, some similar geographic measure could be used. If blocks tend to be too large because of physical size or high population density, the block can be subsampled further so that only a "block face" is used. This maintains the contiguity of the units while reducing the number of units to be matched in any one area.

Another problem the block sample can be helpful in resolving is the allocation of the undercount. The sample blocks can serve as the units of analysis in regressions or log linear models which define how the undercount is distributed in relation to other census variables. These models can then be used to predict the expected undercount for blocks not in the sample. These techniques are described more fully in chapter 6.

A final advantage of the use of the block sample is the ability to take the census listing and match back to the *P*-sample listing. That is, the *E*-sample would be drawn as a set of blocks that coincides with the *P*-sample blocks. This would facilitate *E*-sample matching, allowing the *E*-sample interview to be done in conjunction with *P*-sample followup work, and the high correlation between the *P*- and *E*-samples would lead to smaller variance estimates on the net undercount. A final note on the use of the block sample is that it works best with Procedure A in combination with Definition II. Other designs that are less compact may be more appropriate for other combinations of Procedures and Definitions.

2.42 <u>Use of a survey already "in place"</u>.--Many countries may have an existing survey that can be used for conducting a post enumeration survey. The survey may be multipurpose in nature or may be designed to measure particular types of characteristics. The survey must comprise a probability sample of the entire population of the country. Very often, for purposes of measuring census coverage error, the target population for the survey will have to be supplemented in order for it to correspond to the target population of the census (for example, a sample of the institutionalized population may have to be added). Although there are advantages to using this type of survey as a PES, if a choice exists between using an existing survey or a specially designed survey, the following disadvantages will have to be considered:

(1) The existing survey might have to be supplemented to such an extent that cost benefits would favor a specially designed survey.

(2) The existing survey having been designed for measuring characteristics other than coverage errors could yield coverage estimates with poor accuracy.

(3) A high noninterview rate in the existing survey could bias the coverage error estimates. Very often these types of surveys are on a tight time schedule and are therefore willing "to live with" relatively high noninterview rates. These types of nonresponses are very likely to be missed in the census.

(4) Coverage of the population will probably be weaker with a specially designed survey and could be disproportionately weaker for the very groups in the population that are likely to have census coverage problems.

(5) Very often evaluation planners who use a survey designed for another purpose will have very little latitude in changing procedures, monitoring and controlling the interviewing and processing, and changing the questionnaire to any degree.

(6) The unit selected at the last stage of selection may not be conducive to implementing an efficient erroneous enumeration sample; e.g., it may be a non-compact cluster.

(7) The existing survey EA's were likely designed on the basis of geography defined in a prior census. If the EA's have been redefined in the current census, all the addresses for the survey will have to be assigned a current EA code. These codes are necessary for the matching operation. This assignment is often very difficult. Extensive work in the field that includes having the enumerators draw map locations may have to be done to facilitate the matching operation.

The use of an existing survey for conducting a PES, on the other hand, allows PES planners to use their efforts and resources for other important matters rather than design work which would be required for a special survey. In addition, assuming that the existing survey has been operating with a staff working independently of the census, independence from the census is more likely.

2.43 Erroneous enumeration (*E*-sample) design.--Irrespective of the definition used, a sample of census enumerations (i.e., the *E*-sample) will have to be interviewed to determine if census enumerated "persons" were erroneously enumerated. The sample of census enumerations may be selected independently of the PES sample or they can comprise the same segments that were selected at the last stage of selection for the PES (an overlapped sample). The main disadvantage of an independent sample is cost. The sample of census enumerations does not have to be large as erroneous enumerations occur on a relatively infrequent basis in most countries and thus contribute relatively little to the variance of the net coverage error estimate. In fact, certain countries have omitted this part of the post enumeration survey operation (e.g., Korea and Paraguay).

The major benefits of overlapping the post enumeration survey and census erroneous enumeration segments are twofold: improved precision resulting from the correlation of misses and erroneous enumerations and reduced cost. With regard to precision, a positive correlation between census misses and erroneous enumerations will result in a reduction of variance on the net coverage error estimate. Most of the households in the PES and the census will match exactly in an overlapped sample. Further, for those cases that are non-movers, an erroneous enumeration interview will not be needed; thus followup will only have to be done on a relatively small number of cases.

Experience has shown (1950 U.S. Census PES) that PES enumerators should not be given census information needed for determining census erroneous enumerations. This information could prove a detriment in maintaining independence between the PES and the census. If separate enumerators are not available to do the census erroneous enumeration interviewing at the same time the PES interviewing is done, one would have to wait until the initial PES matching was completed in order to conduct the erroneous enumeration interview. The time lapse could cause a significant memory bias in the resultant estimates, and a follow-up of out-movers would be difficult.

If the PES matching operation is restricted to designated enumeration areas which were reported as locations where the person might have been living (or staying), a separate operation will be required to measure the number of census enumerated locations assigned to the wrong EA. If a location is assigned to the wrong EA, it will be classified incorrectly as a miss in the PES matching operation. The methodology involved in estimating these "erroneous enumerations" consists of "map spotting" the location in the field on existing maps and reconciling any differences (independently) in the PES processing office. The "map spotting" can take place when the interview is conducted for the census erroneous enumeration sample.

2.5 PES *questionnaire design for alternative procedures and definitions*

Associated with each procedure and definition is a unique questionnaire--that is, a different questionnaire has to be designed for each of the optional Procedures and Definitions given in earlier sections. The following are some suggestions that should be considered in designing a post enumeration survey questionnaire.

(1) Procedure A - Definition I--The questionnaire must ask about all persons who resided at the sample address on census day and all locations where those persons might have been enumerated in the census.

(2) Procedure A - Definition II--The questionnaire asks about all persons who resided at the sample address on census day, and assumes they were not listed at any other address.

(3) Procedure B - Definition I--The questionnaire obtains a listing of all persons who currently live at the sample address and obtains all addresses for each person where the person might have lived on census day.

(4) Procedure B - Definition II--The questionnaire obtains a listing of all persons who currently live at the sample address and establishes a single address for each person where they were to have been correctly enumerated in the census.

(5) Procedure C - Definition I--The questionnaire obtains a listing of all persons who currently live at the sample address and all possible locations of these persons on census day (as in Procedure B), as well as a listing of persons who resided at the sample address on census day (but were not resident at the time of the PES) and the locations where these persons might have been enumerated in the census (as in Procedure A).

(6) Procedure C - Definition II--The questionnaire obtains a listing of all persons currently living at the sample address and a single address where each person listed was to have been correctly enumerated in the census, plus a listing of persons who resided at the sample address on census day.

An illustrative post enumeration survey questionnaire for the *P*-sample survey showing variations for each of the combinations of Procedures and Definitions described above is provided in figure 2-3. The core questions shown in sections I, II, and III will be the same for all combinations of Procedures and Definitions. Section IV illustrates prototype questions for each combination of Procedures and Definitions. It should be noted that figure 2-3 does not include items necessary for content error evaluation (see chapter 3). These items would have to be added for content error evaluation.

The choice of Definition (I or II) will also determine the content of the questionnaire administered to the erroneous enumeration sample (i.e., *E*-sample) households as follows:

(1) Erroneous Enumeration Sample - Definition I--This is a multiple purpose questionnaire

designed to be used to estimate duplication in the census, geographic coding errors, and erroneous enumerations excluding the determination of the correct address for enumeration.

(2) Erroneous Enumeration Sample - Definition II--This is the same as for Definition I, but this questionnaire must also determine the one correct address where each respondent should have been enumerated if the sample address is not correct.

An illustrative erroneous enumeration survey questionnaire showing the items of information required for these procedures is provided in figure 2-4. The same questionnaire is used for both Definition I and Definition II, however, the questions used may vary according to the degree of followup.

2.51 Alternative addresses.--Definition I and Procedure B will require the collection of alternative addresses. Very often, these alternative addresses will be incomplete and it will be correspondingly difficult to locate the enumeration area within which the matching search should be conducted. Certain items added to the questionnaire are likely to facilitate this search. These items include questions on names of nearby villages, estates, etc., in rural areas or barrios and neighbors if in urban areas; a description of the location that includes major roads, landmarks, rivers, creeks, etc; the names of the "census day" occupants at that location; and the names of nearby or next door neighbors.

2.52 Probing questions.--Very often probing questions are useful for uncovering persons that the respondent unintentionally left off the household roster. This is especially important for babies, lodgers, relatives, persons travelling or away on business, and so forth. An example of a probing question which might be used to ensure a complete household listing is shown in figure 2-3 under section IV as item 14.

2.53 Interview control.--A record should be kept on the questionnaire of contacts that were attempted with the household (both unsuccessful and successful). A record should be made of both occupied and vacant units. For occupied units, the interview status (complete interview, refusal, not at home, etc.) should be given. Section II of the questionnaire shown in figure 2-3 gives an example of an interview control record.

2.54 Census matching information.--It also may be useful to include space on the questionnaire to include the actual census day roster. This listing of persons should be completed after the actual PES interview at the time of the matching operation. Match status may then be indicated on the actual questionnaire. This procedure facilitates the conversion of data to machine readable form.

2.6 *Matching*

In this manual, procedures are developed for conducting a clerical matching operation. While selecting, training, and supervising a large staff of matching clerks can be difficult, the alternative would be a computer matching operation which may not be feasible in many countries. A computer match requires at least some part of the person's name and a description of the location where he/she was enumerated in the census to be read or keypunched on computer tape. However, the match would be greatly facilitated if an identification number associated with each person, such as a social security or other national identification number, is collected in the census. The major advantage of a computer match would be that the matching could be undertaken in an objective (i.e., non-judgemental) manner utilizing matching algorithms which have a theoretical mathematical basis. Fellegi and Sunter (1969) and Tepping (1968)

Figure 2-3. ILLUSTRATIVE POST ENUMERATION SURVEY QUESTIONNAIRE

POST ENUMERATION SURVEY	CONFIDENTIAL — This inquiry is required by law. The information is accorded confidential treatment and cannot be used for taxation, investigation, or regulation.	
Section I. IDENTIFICATION		
A. Stratum	D. Barrio/village	G. Sample Sector code
B. Province	E. Place	H. Sector letter
C. District	F. EA number	I. Serial number
J. Address or location		

Section II. INTERVIEW RECORD		
K. INTERVIEW STATUS ☐ Unit now occupied 1 Interview completed Interview not completed-- 2 Refusal 3 Partial refusal 4 Not at home 5 Other *(Specify)*______ ☐ Unit now vacant	L. PRINCIPAL RESPONDENT 1 Person on line ____ 2 Neighbor 3 Manager 4 Other *(Specify)* ______	Notes:
Interviewer's name	Date of interview	

Section III. LIVING QUARTERS		
H-1. Type of living quarters 1 Conventional 2 Improvised HU *(makeshift shelter of waste materials, barn, cave, warehouse, etc.)* 3 Mobile HU *(tent, boat, wagon, etc.)* 4 Collective quarters 5 None	H-3. Walls—main construction material of outer walls of this house (building) 1 Stone, cement, stucco, brick 2 Metal 3 Wood 4 Bamboo, leaves, reed, mud 5 Other materials	H-5. Is this unit owned by someone living in it or is it rented? OCCUPIED 1 Owned or being bought 2 Rented for cash 3 Rent free or other arrangment 4 VACANT
H-2. Number of housing units in this house (building) 1 1 unit, detached 2 1 unit, attached 3 2 to 4 units 4 5 to 9 units 5 10 or more units	H-4. How many rooms are in this housing unit? ______rooms	VACANT UNITS H-6. Vacancy status 1 For rent or sale 2 Not for rent or sale H-7. Condition of vacant unit 1 Habitable for year-round use 2 Habitable for seasonal use 3 Not habitable

Figure 2-3. ILLUSTRATIVE POST ENUMERATION SURVEY QUESTIONNAIRE--Continued

PROCEDURE A: Definition I

Section IV. CEN

Line number	What are the names of all persons who lived (or stayed) here on July 1, 1980? Be sure to include babies, elderly persons, and persons who may have been away on vacation or holiday, or may have been in the hospital. *List in this order:* *Head* *Spouse of head* *Never married children of head or spouse (by age)* *Ever married children of head or spouse and their families (by age)* *Other relatives* *Nonrelatives*	What was ...'s relationship to the head of the household? *Interpret categories 3 to 7 to mean relationship to head or spouse.* 1 Head; 2 Spouse of head; 3 Son/daughter; 4 Spouse of son/daughter; 5 Grandchild or great-grandchild; 6 Parent; 7 Other relative; 8 Nonrelative	Sex	What was ...'s age as of last birth-day?	Is ... still residing in this housing unit (at this address)? *If YES, go to item 8.*	What is ...'s current address? *Enter complete address (house number, street, city, village, district, province) or description of location.*
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1		1 Head 3 S/d 5 Gd 7 Other 2 Sp 4 Sp of s/d 6 Par 8 Nonrel	1 M 2 F		1 Y 2 N	

PROCEDURE A: Definition II -- *Includes items 1 to 7, 13, and 14.*

PROCEDURE B: Definition I

Section IV. CU

Line number	What are the names of all persons currently living (or staying) in this housing unit? Be sure to include all babies, elderly persons, and persons who may be away on vacation or holiday, or are in the hospital. *List in this order:* *Head* *Spouse of head* *Never married children of head or spouse (by age)* *Ever married children of head or spouse and their families (by age)* *Other relatives* *Nonrelatives*	What is ...'s relationship to the head of the household? (Categories and instructions the same as in Procedure A.)	Sex	What was ...'s age as of last birth-day?	Where did ... reside (live) on July 1? *If HERE, go to item 8.*	*Enter complete address (house number, street, city, village, district, province) or description of location.* *Go to item 10.*
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1			1 M 2 F		1 Here 2 Else-where	

PROCEDURE B: Definition II -- *Includes items 1 to 9, 13, and 14.*

13. *If listing is continued on additional questionnaire, mark X in this circle* ──► ◯	14. I have listed _____ persons. *Name each person listed.* Have I missed anyone? *If YES, add person to item 2.*

Figure 2-3. ILLUSTRATIVE POST ENUMERATION SURVEY QUESTIONNAIRE--Continued

SUS DAY RESIDENTS

Are there any other places ... could have been living (or staying) on July 1? *If NO, end interview.* (8)	*Enter complete address (house number, street, city, village, district, province) or description of location.* (9)	What is the name of the head of the household or another occupant at that address? (10)	Do you know the names of two next-door neighbors of that address? (11)	*Enter names of neighbors; make sure you get as complete names as possible.* (12)
1 Y 2 N			1 Y 2 N	

RRENT RESIDENTS

Are there any other places ... could have been living (or staying) on July 1? *If NO, end interview.* (8)	*Enter complete address (house number, street, city, village, district, province) or description of location.* (9)	What is the name of the head of the household or another occupant at that address? (10)	Do you know the names of two next-door neighbors of that address? (11)	*Enter names of neighbors; make sure you get as complete names as possible.* (12)
1 Y 2 N			1 Y 2 N	

Figure 2-3. ILLUSTRATIVE POST ENUMERATION SURVEY QUESTIONNAIRE--Continued

PROCEDURE C: Definition I

Section IV. CU

Line number	What are the names of all persons currently living (or staying) in this housing unit? Be sure to include all babies, elderly persons, and persons who may be away on vacation, holiday, or business, or are in the hospital. *List in this order:* *Head* *Spouse of head* *Never married children of head or spouse (by age)* *Ever married children of head or spouse and their families (by age)* *Other relatives* *Nonrelatives*	What is ...'s relationship to the head of the household? *Interpret categories 3 to 7 to mean relationship to head or spouse.* 1 Head 2 Spouse of head 3 Son/daughter 4 Spouse of son/daughter 5 Grandchild or great-grandchild 6 Parent 7 Other relative 8 Nonrelative	Sex	What was ...'s age as of last birth-day?	Where did ... reside (live) on July 1? *If HERE, go to item 8.*	*Enter complete address (house number, street, city, village, district, province) or description of location.*
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1						

PROCEDURE C: Definition I--Continued

Section V. CEN

What are the names of any other persons who lived (or stayed) in this housing unit on July 1?	What was ...'s relationship to *(name of head listed in item 2)*? (Categories and instructions the same as in item 3.)	Sex	What was ...'s age as of last birth-day?	Are there any other places ... could have been living (or staying) on July 1? *If NO, end interview.*	*Enter complete address (house number, street, city, village, district, province) or description of location.*
(13)	(14)	(15)	(16)	(17)	(18)

22. *If listing is continued on additional questionnaire, mark X in this circle* → ○	23. I have listed ____ persons. *Name each person listed in* Have I missed anyone who is currently living here or could here on July 1? *If the answer is YES, add the persons to*

PROCEDURE C: Definition II --*Includes items 1 to 7, 10 to 16, 22, and 23.*

Figure 2-3. ILLUSTRATIVE POST ENUMERATION SURVEY QUESTIONNAIRE--Continued

RRENT RESIDENTS				
Are there any other places ... could have been living (or staying) on July 1? *If NO, end interview for person.*	*Enter complete address (house number, street, city, village, district, province) or description of location .*	What is the name of the head of the household or of another occupant at that address?	Do you know the names of two next-door neighbors of that address?	*Enter name of neighbors; make sure you get as complete names as possible.*
(8)	(9)	(10)	(11)	(12)

SUS RESIDENTS		
What is the name of the head of the household or of another occupant at that address?	Do you know the names of two next-door neighbors at that address?	*Enter names of neighbors; make sure you get as complete names as possible.*
(19)	(20)	(21)

items 2 and 13. have been living *the appropriate item.*	

provide models that can be employed for countries that have the capability to do a computer match.

2.61 Matching procedure.--The following are steps in matching the results of the PES to the census:

(1) Determine the enumeration area (or areas) to be searched. This will require clerks specially trained in the geography of the country and relatively good maps with which to work.

(2) Search for the household(s) listed on the PES form in the specified enumeration area (or areas). Very often the clerk can look for a group of names at the same time. Allowance should be made for misspelling's and misrecordings. If a listing sheet of names and addresses is used in the census, it may be more efficient to initially find the household on this sheet, and obtain a cross referenced serial number that would help to locate the desired questionnaire in a more efficient manner.

(3) Determine which persons listed on the PES form are also listed on the census questionnaire. Possible matches should be examined very carefully, paying strict adherence to the matching rules.

(4) Cases which cannot be classified on the basis of the information available should be referred to the field followup operation to collect additional information to permit proper classification of match status.

2.62 Matching rules.--Detailed rules for matching will have to be prepared. In developing these rules, one has to consider the possibility of making erroneous matches and erroneous non-matches. Erroneous matches are defined as cases which are classified as matches when in fact the PES case was not actually enumerated in the census.

Erroneous non-matches are defined as cases classified as non-matches which in fact do correspond to a case enumerated in the census. If exact agreement of characteristics is needed to establish a match, an excessive number of erroneous non-matches is likely to occur. As rules are relaxed to allow more matches (and fewer non-matches), an increasing number of erroneous matches will occur. The objective of evaluation planners should be to design a system which minimizes net error (the difference between erroneous matches and erroneous non-matches). The objective of the matching procedure is to determine the number of matches (as described later in this chapter); this estimate will be accurate if the net matching error is equal to zero.

Matching rules specify the characteristics (e.g., name, sex, age, etc.) by which persons and households enumerated in the census and PES are to be matched. Tolerance ranges within which records must agree, should also be defined. These tolerances will allow for a limited degree of misreporting in either the PES or the census. Tolerances can vary according to characteristics; for example no tolerance may be allowed for sex differences, but relatively large tolerances may be allowed for age.

Tolerances might also vary according to the situation; for example, if it is certain that the desired household has been identified, it may be useful to relax tolerance standards for the individuals within the household. By varying the number of characteristics and/or the tolerance limits one can vary the size and sign of the net matching error. As the number of characteristics used for criteria in performing the matching increases, the number of erroneous non-matches will increase and the number of erroneous matches will decrease. Conversely, as tolerances for the matching of characteristics are increased, the number of erroneous non-matches will decrease and the number of erroneous matches will increase.

One implication of this is that it is possible for the net matching error to be

Figure 2-4. ILLUSTRATIVE ERRONEOUS ENUMERATION SURVEY QUESTIONNAIRE

ERRONEOUS ENUMERATION SURVEY	CONFIDENTIAL — This inquiry is required by law. The information is accorded confidential treatment and cannot be used for taxation, investigation, or regulation.

Section I. IDENTIFICATION

A. Stratum	D. Barrio/village	G. Sample Sector code
B. Province	E. Place	H. Sector letter
C. District	F. EA Number	I. Serial number

J. Address or location ______________________

____

____

____

Section II. CENSUS STATUS

K. CENSUS STATUS

☐ Occupied

☐ Vacant

☐ Unknown

☐ Imputation case

Section III. SURVEY STATUS

L. SURVEY STATUS

☐ Unit not found
☐ Unit no longer exists
} *(Explain in NOTE section)*

☐ Unit vacant

☐ Unit occupied → ☐ Interview completed
☐ Refusal

RESPONDENT

M. PRINCIPAL RESPONDENT/SOURCE OF INFORMATION

☐ Occupant

☐ Neighbor

☐ Manager

☐ Other *(Specify)*______________________

Respondent's name

____

Respondent's address or location

____

____

Notes

Interviewer's name	Date of interview

Figure 2-4. ILLUSTRATIVE ERRONEOUS ENUMERATION SURVEY QUESTIONNAIRE--Continued

Section IV. CENSUS INFORMATION *Transcribe items 2 to 6 from census questionnaire for each person.*								
Line number	Name	Relationship 1 Head 2 Spouse 3 Son/daughter 4 Spouse of son/daughter 5 Grandchild or great-grandchild 6 Parent 7 Other relative 8 Nonrelative	Sex	Age	Marital Status 1 Married 2 Widowed 3 Divorced 4 Separated 5 Never married	DO YOU KNOW ...? *If unit now vacant, go to Section VI.* *If answer is NO for a person, complete Section VI after completing interview for the other persons.*	DID ... LIVE OR STAY HERE (IN THIS HOUSING UNIT) ON JULY 1? *If yes, go to item 13.*	DO YOU KNOW WHERE ... LIVED OR STAYED ON JULY 1? *If YES, go to item 12.*
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1		1 Head 5 Gd 2 Sp 6 Par 3 S/d 7 Oth 4 Sp of s/d 8 Non	1 M 2 F		1 Mar 4 Sep 2 Wid 5 NM 3 Div	1 Y 2 N	1 Y 2 N	1 Y 2 N

Section VI. INTERVIEW AT OTHER ADDRESSES						
DO YOU KNOW ...? *If NO, go to Section VII.*	DID ... LIVE OR STAY AT *(sample HU address)* ON JULY 1? *If YES, go to item 18.*	DO YOU KNOW WHERE ... LIVED OR STAYED ON JULY 1? *If YES, go to item 20.*	DO YOU KNOW ...'s CURRENT ADDRESS (PLACE OF RESIDENCE)? *If NO, go to Section VII.*	*Enter complete address (house number, street, city, village, district, province) or description of location.* *When interview completed for all persons in item 2, go to current address to get information needed.* *End interview for this person.*	*Enter complete address (house number, street, city, village, district, province) or description of location.*	IS THERE ANY OTHER PLACE ... COULD HAVE LIVED OR STAYED ON JULY 1? *If NO, go to Section VII.*
(15)	(16)	(17)	(18)	(19)	(20)	(21)
1 Y 2 N	1 Y 2 N	1 Y 2 N	1 Y 2 N			1 Y 2 N

Figure 2-4. ILLUSTRATIVE ERRONEOUS ENUMERATION SURVEY QUESTIONNAIRE--Continued

Section V. INTERVIEW AT SAMPLE ADDRESS				
DO YOU KNOW ...'s CURRENT ADDRESS (PLACE OF RESIDENCE)? *If NO, go to Section VI.* (10)	*Enter complete address (house number, street, city, village, district, province) or description of location.* *When interview is completed for all persons, complete Section VI.* (11)	*Enter complete address (house number, street, city, village, district, province) or description of location.* (12)	IS THERE ANY OTHER PLACE ... COULD HAVE LIVED OR STAYED ON JULY 1? *If NO, end interview.* (13)	*Enter complete address (house number, street, city, village, district, province) or description of location.* (14)
1 Y 2 N			1 Y 2 N	

	Section VII. PRINCIPAL RESPONDENT	Section VIII. INTERVIEWER CHECK	
Enter complete address (house number, street, city, village, district, province) or description of location. (22)	Principal respondent for Section VI. (23)	*Apply census day residency rule for sample housing unit and all other locations. Mark the address or location you feel meets the census residency criteria. Specify reason.*	
	1 Neighbor *(Specify)* ________ 2 Manager *(Specify)* ________ 3 Other *(Specify)* ________ 4 Unable to find a knowledgeable respondent	1 Sample address 2 Item 11 or 19 3 Item 12 or 20 4 Item 14 or 22	Reason ________

small, but the gross error (erroneous matches plus erroneous non-matches) to be large. Hence, designing a system that depends on such trade-offs could be very risky. If the actual matching operation is somewhat less efficient than was anticipated, the net error could be disastrously large. In addition, even if the net matching error is small, the variance of this estimate could be large if the two components of coverage error are large. A "safer" approach would be to try to minimize both the net matching error and the gross error to the extent possible.

The matching process is perhaps the most expensive and difficult part of the PES operation. Some form of the following three-tiered matching system, which has proven to be relatively efficient in previous experience, is recommended:

(1) The first "tier" of matching would have clerks classify definite matches only. Very tight tolerances would be used keeping erroneous matches to an absolute minimum. Clerks would not need extensive training in order to process these "easy" cases.

(2) The second "tier" consists of a specially trained set of clerks who use more complex rules to reclassify the "remaining cases" as definite matches, definite non-matches, or "status unknown" cases (that is, "possible" matches).

(3) The third "tier" consists of supervisors and professionals who work together to resolve the "status unknown" cases. Any information which can be found on the problem case can be used at this stage. Out of this operation will come definite matches, definite non-matches, and cases that need to be followed up in the field to obtain additional information (the expectation is that these cases should be few in number; otherwise another expensive field operation would be required).

"Optimum" matching rules, which would attempt to minimize erroneous non-matches and erroneous matches, would be used for the second tier and "problem cases". The tiered approach has been found to work quite well in the United States and in several developing countries (e.g., Bangladesh and Egypt). One of the strengths of this approach is that it permits relatively strict operational and quality control procedures to be implemented.

Matching characteristics, to a great extent, depend upon the country where the study is being done. However, the following <u>illustration</u> of suggested rules for determining matches (and non-matches) may prove useful as a start in designing a matching operation.

(1) <u>Rules for first tier definite matches.</u>

The case will be a match if

(a) place of residence is in the same or an adjacent EA and agrees exactly except for a rearrangement of street, building, or dwelling unit numbers or a similar sounding name. Designations also may be different, for example, road for avenue or street for avenue, etc.;

(b) family name (surname) agrees except for minor spelling differences which do not change the sound. For a common name, given and middle names of the subject must agree except for minor differences;

(c) relationship is not contradictory;

(d) sex agrees exactly; and

(e) age is within -1 to +2 years for persons under 20, -2 to +3 years for ages between 20 and 40, and -3 to +4 years for persons over 40.

After this matching operation has been completed, a set of clerks can examine the remaining uncertain cases with the following set of "second-tier" guidelines to determine a match status.

(2) <u>Rules for second tier matches.</u>

The case will be a match if the same conditions for a definite match in the first tier matching operation are satisfied except that

(a) there is a contradiction in <u>one</u> of the following--given (or first) name, relationship, or sex;

(b) address information is contradictory; for example, address information may be missing from either the census or the PES; or

(c) large age differences (as determined from date of birth) are allowed, such as the following:

Age	Tolerance
Under 10 years	± 1 year
10 to 19 years	± 2 years
20 to 39 years	± 6 years
40 to 59 years	± 8 years
Over 60 years	±10 years

After the cases have been reviewed by the second tier and further cases specified as matches, all remaining cases will either be definite non-matches or cases for which match status still cannot be determined. Indeterminate cases and, if resources permit, non-matched cases, will go on to the third tier where rules are much more subjective in nature. Both match and non-match determinations will result out of this operation. (Note: as time progresses, and the clerical staff gains experience it may be feasible to combine the tier 1 and tier 2 operations.)

It would be desirable for planners to experiment with various combinations of characteristics and tolerances in one or more pretests. Assuming that all items have relatively low response errors associated with them, the combination of characteristics and tolerances yielding the smallest net error and gross error would be the most desirable. The decision on matching rules assumes knowledge of the probabilities of erroneous matches and erroneous non-matches. This information might be obtained from a pretest by collecting sufficient matching information from respondents so that the resolution of questionable cases can be made, although it is likely to be infeasible to obtain this "additional" information in the regular PES. Another option might be to obtain this additional information for a subsample of the regular PES; in this case, initial matching could be confined to this subsample such that matching rules for the remainder of the PES may be established.

One method which may be useful for obtaining optimal tolerances <u>for a particular characteristic</u> would be to start with zero tolerance and gradually increase the limits. At each step record the number of erroneous matches and erroneous non-matches that correspond to this characteristic. Tolerance limits would ultimately be set at the level for which the increase in erroneous matches is greater than the decrease in erroneous non-matches; that is, the gross matching error is minimized.

In order to determine the optimal number of characteristics to use, the procedure is a bit more complicated. A suggested procedure for when tolerance limits are fixed is as follows:

(1) Select a characteristic that has the lowest probability of an erroneous match (this will usually be a characteristic with the smallest net error). Calculate the net error for this characteristic.

(2) Select the next characteristic to be used by choosing the one that will yield the greatest reduction in <u>net error</u>.

(3) Continue on with succeeding characteristics until the sign of the net error is changed.

A worksheet can be developed that may be used to list all relevant information needed to make a decision on the optimal number of characteristics to use for matching. The estimation of the relevant probabilities is discussed in the section on estimation that discusses matching bias (section 2.84).

It is desirable for matching rules to be explicitly stated and documented so that persons who do the matching do not apply different criteria. In some cases, subjective matching may give better results; but, over a large number of matching decisions, an operation that uses explicitly stated matching rules will be much more controllable and statistically defensible, as it will likely yield estimates with smaller net and gross matching errors.

2.63 One way versus two way match.--If the census erroneous enumeration sample (*E*-sample) is selected independently of the post enumeration survey, the match to measure census underenumeration will be strictly in one direction (from the post enumeration survey to the census records). However, if the census erroneous enumeration sample is overlapped with the post enumeration survey, a match in both directions (two way match) could result in substantially less field work for the census erroneous enumeration survey. This is true under both Definitions I and II. Census erroneous enumeration interviews will only have to be made for persons who are listed on the census questionnaire but are not on the PES questionnaire (under the assumption that the address is the same for the two households). For persons who are on both the PES form and the census form, it is illogical for PES questions to be repeated in the census erroneous enumeration survey (which would likely result in a repeat of the PES answers).

2.64 Followup interview to obtain additional information.--After the matching has been completed, there will likely be some cases for which the matching status is not ascertainable. The available information may be so vague or incomplete that it is impossible to determine if the person was enumerated in the census with any reasonable degree of certainty. For many of these cases, a followup interview to collect additional information may resolve the matter. For cases which remain uncertain even after the field followup, a noninterview adjustment will have to be introduced into the estimation procedures (see section 2.82).

2.7 *Special problems*

There are certain problems one needs to be aware of and take steps to deal with in order not to bias the coverage estimates. The most serious problems are described below.

2.71 Followup of movers.--A major problem with both Procedure A and the census erroneous enumeration sample is the enumeration of out-movers. The followup of out-movers is expensive, time consuming, and the results are often less than successful; however, it is necessary to undertake such an operation in order to obtain "unbiased" results (assuming no response error occurs). If followup is not conducted, information on out-movers will have to be collected from proxy respondents and an unknown and possibly significant bias can be a consequence. For the *E*-sample, the bias can be particularly serious for Definition II used in conjunction with Procedure B. A typical pattern of response error often made with Procedure B is for the respondent to report residing at the current sample address on census day when in fact the person was an in-mover. If an overlapped census erroneous enumeration sample is used, the in-movers would provide a proxy response for the census out-movers if no followup is done on out-movers; thus, the out-movers would probably be classified as not living at the current address on census day, resulting in a census erroneous enumeration. In this situation, the current residents will be missed in the census (erroneously) and the census day residents of the PES sample address also will have been erroneously enumerated (since they were not followed up to obtain an interview). This would, however, depict the correct situation if the current residents were enumerated at their previous address; if not, a bias will result.

2.72 Assignment of current EA codes to the PES.--A problem in matching can occur if a current or ongoing survey is used as a vehicle to estimate coverage error, particularly if the survey was designed on the basis of enumeration areas from a previous census which had different enumeration area definitions. Therefore, all current survey addresses must be assigned enumeration area codes that correspond to the current census. This can be a time-consuming task, and if done improperly can result in the coverage estimates being severely biased. It is very important for the current survey enumerators to prepare high-quality location maps for the sample addresses so that the assignment of EA codes in the PES can be accurately undertaken.

2.8 *Estimation*

Some earlier post enumeration surveys were designed to give single system estimates of coverage error; for example, the PES for the 1950 Census of the United States. This type of PES had as its primary objective superior coverage in the survey than in the census. More intensive canvassing procedures, more experienced interviewers, closer supervision, and more extensive probing are employed in an effort to produce better population coverage. Efforts are made to include all the persons who were missed in the census as well as all who should have been enumerated in the census and who actually were. Census results may be used in a reconciliation operation in an effort to improve the quality of coverage in the PES; thus less attention is given to maintaining independence with the census. The following estimator is applicable to this type of PES:

Let:

N' = the estimate of the total population

U_2' = the estimate of the number of people that were in the PES but not in the census (obtained by matching the PES to the census)

N_1 = the census count

The estimate of the total population is given by:

$$N' = N_1 + U_2'$$

and the census miss rate by:

$$R_m = \frac{U_2'}{N'}$$

These estimators are conceptually biased because they do not take into account the category "missed in both the PES and the census". An evaluation of the 1950 U.S. census PES results verified that this bias was serious. In the evaluation, comparisons were made between the PES and estimates derived on the basis of demographic analysis. For important subgroups of the population, the PES estimates badly understated the magnitude of the undercount. Apparently, persons enumerated in the census were much easier to locate and enumerate in the PES than persons missed in the census. Some persons missed in the census were not reported in the PES for the same reasons that they were not reported in the census, such as: deliberate concealment, ambiguity of residence rules for persons with little attachment to a given household, and isolated or hidden dwelling units.

Thus, a "best quality" type of PES may reduce bias due to poor enumerators or careless canvassing procedures, but does little to reduce the correlation bias that arises from the "nature" of the population that is to be enumerated. In addition, the reconciliation operation which is a feature of the "best quality" PES approach can actually increase correlation bias, since it eliminates PES errors which erroneously report persons who were not living at a given address on the census date and were, therefore, not enumerated at that location in the census. The reconciliation operation does not properly account for errors made in the PES of omitting persons who should have been enumerated at an address in the census, but were not.

Furthermore, there is substantial evidence that even though a survey may devote considerable attention to population coverage, it will still not achieve as good coverage as a moderately well executed census. Thus, it is currently felt by many PES practitioners that the "best quality" type of PES is doomed to less than successful outcome.

Another PES approach that has been used in some countries is the following:

Let:

N_2 = the PES estimate of the total population

U_1' = the estimate of the persons enumerated by the census and missed by the PES

At the time of the PES interview, the interviewer is given the census roster to be used to improve the PES roster. Estimates of the total population and the census miss rate can be derived from this information as follows:

$$N' = N_2 + U_1'$$

$$R_m = \frac{U_1'}{N'}$$

This method requires no matching operation in the central statistical office. However, it has several weaknesses. Not only is it subject to the bias of omitting persons missed in both the PES and the census, but it also is potentially subject to increased variance in comparison with the estimators described above because of potential correlations between U_1' and N_2.

2.81 Dual system estimation.--The general model for dual system estimation was described in section 2.1 of this chapter. Subsequent sections described the data collection process and the use of the *P*- and *E*-samples to control problems of missing and erroneous data. Dual-system estimates of the size of the total population may be obtained from the survey results as:

$$\hat{N}_T = \frac{N_p\,(N_c - G - E - D - I)}{M} \qquad (2.1)$$

where

N_p = the weighted total of persons in the *P*-sample

N_c = the total number of persons enumerated in the census

E = the weighted total of persons erroneously enumerated in the census from the *E*-sample (e.g., born after census day

D = the weighted estimate from the *E*-sample of duplication in the census

G = the weighted estimate of persons misassigned geographically in the census from the *E*-sample

$\hat{N}_T$ = the true population size, the value to be estimated

I = the number of persons **enumerated in the census but** having insufficient information for matching

Similarly, the net undercount rate can be estimated as

$$\hat{R}_N = \frac{N_c}{N_p\,(N_c - G - E - D - I)/M} \qquad (2.2)$$

The observed census total divided by the dual system estimate, or

$$\hat{R}_N = \frac{\dfrac{M}{N_p}}{(N_c - G - E - D - I)/N_c} \qquad (2.3)$$

M = The weighted total of persons matched from the *P*-sample to the census (persons enumerated in both the *P*-sample and the census)

The numerator and denominator of $\hat{R}_N$ represent gross error rates from the census. The numerator is the estimate of the coverage rate in the census, and the denominator is the estimate of the correct enumeration rate in the census. Writing these in more familiar terms,

$$(2.4) \quad \hat{R}_N = \frac{1 - \hat{GU}}{1 - \hat{GO}}$$

where

$\hat{GU}$ = the gross undercount rate (uncorrected for errors in the census) and

$\hat{GO}$ = the gross undercount rate, including census errors

Note that if there were no errors in the census (that is, duplicates, geography errors, etc.), the net undercount rate would equal the gross undercount rate. Note also that the ratio $\hat{R}_N$ is usually less than one, denoting a net undercount. However, $\hat{R}_N$ can also be greater than one, indicating a net overcount.

Each of the components in this estimator is obtained as a sample estimate derived from either the P-sample or the E-sample. In developing the estimates, it is desirable to make weighting adjustments to the E-sample to force the E-sample marginal totals for selected characteristics to exactly fit the census totals for these characteristics, since the E-sample is a subsample of the census.

2.82 Treatment of missing data.--In addition to the development of weights and sample estimates, there is the problem of missing data to be dealt with. Because of time constraints, lack of resources, errors in processing, and reluctance of respondents, some crucial data items will be missing. Depending on the volume and particular items that are missing, the problem may be handled either by making weighting adjustments or imputing the missing data. The U.S. Census Bureau used both techniques in the 1980 census PES. For cases in both the P-sample and the E-sample, in which entire households were missed (for example, a complete refusal to be interviewed for all members of the household), the household was dropped from the sample and a weighting adjustment was made to account for its loss. For cases in which only particular data items were missing, e.g., match status in the P-sample for a case which was an otherwise complete interview, the missing data value was imputed. The imputations were performed using the "nearest neighbor" policy. That is, imputation values were chosen from cases that were most similar to the case for which data were missing. Cases were generally linked by age, race, sex, household size, and proximity.

However the missing data problem is treated, caution should be exercised in choosing a method for making data adjustment, since the choice of the method directly affects the estimates of the undercount. This is particularly true if the undercount rate is small relative to the proportion of data that are missing.

2.83 Post stratification of estimates.--Often different subgroups in the population have quite different rates of coverage in the census. If this occurs in the population, it becomes necessary to post-stratify the dual system estimates to avoid mixing the capture probabilities. A national level estimate, for example, can mix a high undercount rate for the young (and often highly mobile) population with a low undercount rate for the older population, leading to a bias in the estimate. Separate estimates for each age group would avoid these problems, and the sum of the dual system estimates across age or other subgroups will yield an unbiased estimate of the national population total.

As a general rule, it is best to form estimates for subgroups which are as homogeneous as possible, and sum the estimates up to the level of aggregation desired. On the other hand, if the subgroups are very small, a ratio bias can occur in the sample estimates. Generally, subgroups should contain at least 50 persons or housing units. The estimates may subsequently be post-stratified by age, sex, geography, or any other variables which are felt to be important.

2.84 Special problems related to estimation.--The dual system estimation method makes certain assumptions with regard to the two sources of data. If these requirements are not met, biases of the dual system estimates of the true number of cases and the completeness rate will result. These biases are of the following type:

(1) Correlation bias. Correlation bias is the tendency of cases included in the census to have a higher probability of inclusion in the PES than cases not included in the census. This can be a particularly serious problem for certain subgroups of the population. Correlation bias occurs due to the nature of coverage error and the data collection systems, that is, the census and the PES. Very often the same persons tend to be missed in both the PES and the census because they are members of population subgroups which are difficult to cover. Bias also will occur because of "communication" between the PES and the census. This includes any interaction between the field staff and procedures of the PES and the census that affect the coverage or omission of persons in either the PES or the census.

(2) Matching bias. As previously mentioned, there are two types of matching errors: erroneous matches and erroneous non-matches. If the expected value of the difference of these errors is zero, they will have no effect on the dual system estimate (that is, they will cancel each other out). However, as each source of error increases, there also will be an increase in the variance of the estimate. Since it is generally not an advisable statistical policy to rely on the mutual cancelling of different types of errors to obtain accurate results, and because of the detrimental effect on the variance of the dual system estimate, it is best to minimize both types of error.

(3) Variance considerations. Because the PES is a sample survey, the dual system estimator will be subject to sampling variance. Sampling variance will be a function of the sample design used, the sample size, and the census undercount rate. As the degree of clustering goes up or the sample size is decreased, the sampling variance will increase. As the undercount rate increases, the sampling variance will increase.

The variance of the dual system estimator also is affected by nonsampling error. These are errors in household counts attributable both to enumerators and respondents, and exist in both the census and the PES. There are also matching errors and other clerical errors which affect the precision of the estimates. The impact of these errors on the dual system estimator is similar to that encountered for the index of inconsistency in the content evaluation portion of this manual (see chapter 3). There may be an increase or decrease in variance depending upon the net direction of the error. The literature on the implications of nonsampling error for dual system estimation is to date inadequate, however, and further research is needed.

2.9 *Suggested tabulations*

The following tabulations should be considered in order to fully understand the nature of the coverage error problem. Not all of these tabulations will be appropriate for all countries, however. It is important then that considerable attention be given to which types of estimates are to be prepared. These tabulations are necessary not only in deciding what form the "official" coverage error estimates will take, but may also be useful in revealing the intrinsic nature of the coverage error problem. Suggested tabulations are:

(1) Population enumerated in the census for relevant demographic/socioeconomic subgroups by relationship to head of household and sex in major geographic areas of the country; especially important are urban-rural areas where the nature of the coverage error problem can be vastly different.

(2) PES estimates for the same categories described above for the census.

(3) Estimates of census imputations, erroneous enumerations, persons in the census assigned to the wrong census enumeration area, census duplicate enumerations and persons whom the PES identified as living at a residence other than the census reported residence on census day.

(4) Nonresponse rates in the PES by cause of the noninterview (refusals, not at homes, etc.)

(5) Dual system estimates of the total population by type of procedure (A, B, or C) if relevant information is collected to obtain all three kinds of estimates. These estimates should be calculated separately for movers and non-movers as well as for the total.

3. *LONGITUDINAL TRACING STUDIES*

Post enumeration surveys have experienced mixed success in the United States. For some groups of the population, the estimates have been relatively good; however, for other groups they have been less satisfactory (e.g., young males aged 18 to 25). A procedure that has been successfully used in other countries (e.g., Canada), the longitudinal tracing study, has considerable promise in evaluating the coverage of subgroups of the population for which a large post enumeration survey correlation bias exists. The procedure is, however, relatively expensive; and thus may be economically feasible to apply only to specific subgroups of the population, with the post enumeration survey (PES) approach providing coverage estimates for the remainder of the population.

3.1 *Coverage technique*

The longitudinal coverage technique attempts to create independent components of the population which collectively represent the population at any point in time. For a given census, the following population components may be identified: persons enumerated in the previous census, persons missed in the previous census, intercensal births, intercensal naturalized citizens, and registered aliens. A sample of these persons is selected shortly after the most recent previous census and is monitored over the intercensal period to obtain demographic or socioeconomic information and up-to-date address information. Additions are made to the sample between censuses of births and immigrants on a regular basis; thus, the sample actually grows in size between censuses. A match is then made to the current census listings to determine whether the sample persons were enumerated. This matching operation could be structured in much the same manner as a post enumeration survey.

Tracing procedures should be used which minimize potential correlation bias since a person's knowledge of the tracing procedure could influence his/her desire to be enumerated in the current census. Current address information may be obtained from such sources as designated contact persons who are determined in the initial interview, post office "mover" records, or other administrative records. Post office "mover" records or other administrative records would seem to be especially promising for purposes of keeping correlation bias to a minimum. This procedure is presently being tested in the United States as a method of estimating coverage error in the 1990 census.

3.2 *Alternative tracing procedures*

The longitudinal coverage technique is based on the proposition that groups of people who are difficult to enumerate on census day are easier to enumerate, or to include in a sample, some years before the census when they are members of a population segment which tends to be covered more completely in censuses. However, correlation bias may nevertheless arise when those people who were traced successfully are more likely to be counted in the census than those who were not traceable. The manner in which the tracing procedure is conducted is an important determinant of the potential correlation bias that can occur.

The following are five alternative tracing procedures which might be considered:

(1) During the intercensal period, periodic tracing is done which involves no personal contact with the subject (for example, only the post office or other groups are contacted, never the individual).

(2) During the intercensal period, the subject is initially contacted in order to obtain basic information needed for future tracing operations, but no further personal contact is made.

(3) During the intercensal period, the subject is contacted initially to collect basic information, and is then contacted periodically to obtain current address.

(4) No tracing is done or contact made during the intercensal period; at the end of the intercensal period, the subject is retrospectively traced.

(5) No tracing is done during the intercensal period; at the end of the intercensal period, the subject is retrospectively traced. The subject is contacted early in the intercensal period to collect basic information that can be used later in the retrospective tracing operation.

In general, operations that involve contacting the subject are very risky. However, if this can be done without increasing correlation bias, there are potential gains in reducing erroneous non-matches. Some considerations that need to be addressed in further research are:

(1) the costs of the various alternatives

(2) the magnitude of correlation bias caused by contacting the subject

(3) the percent of persons lost during the period of alternative tracing procedures

(4) for retrospective tracing, the proportion of persons who are "lost" for various tracing periods

4. USE OF NETWORK (MULTIPLICITY) SAMPLING IN ESTIMATING COVERAGE ERROR

Recently, research has been conducted at the U.S. Bureau of the Census on the use of network (or multiplicity sampling) for the purpose of estimating the undercount. In a typical household survey, where the household is the sampling unit, a person can be reported if he/she is a member of the household that was selected for the survey. In a multiplicity survey, a person can be included in the survey by either being a member of the household that is selected, or by being linked from another household not in sample. This technique has been used in the United States for estimating the incidence of rare diseases in the population, but only recently has been considered for use in estimating census coverage error. Since "events" can be reported at more than one household residence, there are more chances for an "event" to be reported with the corresponding variance thus being lowered. Multiplicity sampling has also been used for purposes of estimating vital events and the completeness of death registration lists.

The "rare event" that one tries to estimate in a census coverage evaluation program is the incidence of persons being missed in the census. Sample household members are asked to report their census day addresses and those of specified other persons (usually specific relatives). These reported census day addresses are then matched to census records to determine if the subjects were missed. Probabilities of including a missed person (and hence sampling weights) are determined by taking account of the number of persons who could report the subject.

The multiplicity counting rule specifies which individuals a household is eligible to report. Usually the sample households report their de jure (resident) household members, parents, children, or siblings who reside at other addresses. The procedure used in weighting the subject is determined by the counting rule adopted. For example, if the subject can be reported at their de jure residence or at their siblings' households, the respondent must report in the survey how many other households exist in which their brothers and sisters reside. The probability of the respondent's selection in the survey is then the probability of selection of the sample household plus the sum of the probabilities of the siblings' residences being selected.

Multiplicity surveys have certain disadvantages when used as vehicles for estimating coverage error, such as:

(1) The household respondent may not report all other persons they can be linked to even if they are members of the household. (A PES also has this deficiency.)

(2) The household respondent can report other persons as members of the "extended" household, but may not know anything more about the persons. In this case, it would be useful to talk to the other subjects directly. (Again, this same problem can and does arise in a PES.)

(3) The respondent purposely omits certain designated relatives or does not give sufficiently detailed or accurate addresses or demographic information for designated relatives.

(4) Single persons with relatives have only one chance to fall into sample, instead of a multiplicity of opportunities. If census misses are most common among the homeless and indigent, these persons could have a zero probability of selection, and thus still pose a major source of coverage error.

On the other hand, some of the advantages of the multiplicity approach are the following:

(1) A large (network) sample can be obtained at relatively low cost. Thus sampling variances can be substantially reduced over a PES or a tracing study.

(2) The potential for a substantial reduction in correlation bias presents itself if subjects who are missed in the census and omitted from their de jure residence are reported by a designated relative.

4.1 *Counting rules*

Prior to implementing a multiplicity survey, research should be conducted for purposes of determining the optimum counting rule. Combinations of any of the following may be used to determine a counting rule:

(1) all residents of the de jure residence

(2) spouses of residents who reside elsewhere

(3) children of residents who reside elsewhere

(4) parents of residents who reside elsewhere

(5) siblings of residents who reside elsewhere

The ability to obtain good matching information (addresses and demographic information) for the above counting rules can, obviously, vary considerably. Research in the United States with small samples has shown that the best matching information is obtained from parents of respondents and the least reliable information is from siblings of respondents. A particular counting rule should not be used if the rate at which data and cases are lost due to poor matching information is usually high. A counting rule which results in proportionately more data lost than the coverage error being estimated should in particular not be used.

4.2 *Estimation*

The estimators for a multiplicity survey are somewhat different than those for a conventional survey. In a conventional household survey, each person with the desired characteristic in question can be reported only by the household in which they reside. However, in the multiplicity survey, each person can be reported by the household in which he/she resides as well as other households in which he/she does not reside, but which happen to fall in the sample. The number of "other" households reporting the subject will depend upon the multiplicity counting rule selected. The total number of households reporting the subject is referred to as their multiplicity.

One will recall from an earlier section of this chapter that the PES dual system estimator for the total population, N_T, is:

$$\hat{N}_T = N_c \cdot \frac{N_p}{M}$$

Assuming that a self-weighting probability sample of m households is selected for the multiplicity survey sample from M households in the population and a particular counting rule has been used, the multiplicity estimator of the total population, N_T, is given by

$$\hat{N}_T = \frac{1}{f} \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{1}{K_j} (r_{j,i} + S_{j,i})$$

where:

$\frac{1}{f}$ = the inverse of the sampling fraction

N_T = the total number of persons in the population

I_i = the i^{th} person in the population $(i = 1,2,\ldots,N)$

H_i = the i^{th} household in the population $(i = 1,2,\ldots,M)$

$$r_{j,i} = \begin{cases} 1 & \text{if } I_i \text{ is a resident of } H_i \\ 0 & \text{otherwise} \end{cases}$$

$$S_{j,i} = \begin{cases} 1 & \text{if } I_i \text{ is not a resident of } H_i\text{, and is reported by } H_i \\ 0 & \text{otherwise} \end{cases}$$

$$K_j = \sum_{j=1}^{M} (r_{j,i} + S_{j,i}) ,$$

the number of households reporting I_i (the multiplicity for I_i)

Thus to estimate K_j for each member reported by H_i, one must obtain a count of the number of other households which contain relatives of the subject.

The multiplicity estimator for the match total M is:

$$\hat{M} = \frac{1}{f} \sum_{i=1}^{m} \sum_{j=1}^{N} (u_{j,i} + V_{j,i})$$

where:

$$u_{j,i} = \begin{cases} 1 & \text{if } I_i \text{ is a resident of } H_i \text{ and is matched to a census record} \\ 0 & \text{otherwise} \end{cases}$$

$$V_{j,i} = \begin{cases} 1 & \text{if } I_i \text{ is not a resident of } H_i\text{, and is reported by } H_i\text{, and is matched to a census record} \\ 0 & \text{otherwise} \end{cases}$$

5. *PROFILES OF INTERNATIONAL COVERAGE EVALUATION STUDIES*

This section presents brief descriptions of census evaluation studies which have been undertaken in selected countries in recent years. This material is presented to provide some indication of the types of coverage evaluation designs which have been attempted in countries under different circumstances, as well as some of the key findings of these evaluation efforts.

The presentation begins with a discussion of recent evaluation work in two developed countries (Canada and the United States), followed by a description of PES experiences in four developing countries (Korea, Paraguay, Bangladesh, and India).

5.1 *Canada*

A longitudinal coverage approach was initially used in Canada in conjunction with the 1961 census on an experimental basis and later as a primary evaluation methodology for the 1966, 1971, and 1976 censuses. The following description is the methodology utilized for the 1976 census.

The target population for this study was all persons who should have been enumerated in the 1976 census. From this population, a sample was chosen for matching purposes. Four separate frames were used to represent the population who should have been enumerated in the census. This population included:

(1) Persons enumerated at their usual place of residence in the 1971 census.

(2) Registered births between census day of the previous census (June 1, 1971) and census day of the current census (May 31, 1976).

(3) Registered immigrants to Canada between June 1, 1971 and May 31, 1976.

(4) Persons not enumerated at their usual place of residence in the previous 1971 census (that is, persons missed in the 1971 census).

A sample size of 33,000 persons was allocated to these four frames, with estimates of undercoverage available at the national level for broad age-sex groups and geographic regions.

The objective of the study was to classify each person in the sample into one of the following four categories:

(1) Enumerated in the current 1976 census

(2) Missed in the 1976 census

(3) Died before June 1, 1976

(4) Emigrated before June 1, 1976

A tracing operation was conducted to assign each person to a place of residence as of the current census day (June 1, 1976). For this operation, a series of traces was performed retrospectively. The following activities were conducted in order of priority.

(1) Each sample person from the 1971 census frame was matched to the 1976 census to see if he/she was enumerated in the current census at his/her 1971 census address.

(2) For those not matching in (1) above, or in the other frames, a match to administrative records was made to obtain current addresses.

(3) Extensive telephoning was done to obtain addresses at the time of the 1976 census.

(4) Fieldwork was done whereby the subject was personally contacted to obtain current census address information.

The results of the study were the following:

Final Status	Unweighted sample results (percent)
enumerated	88.2
missed	2.5
deceased	3.2
emigrated	1.3
trace failures	4.8

5.2 *United States*

Post enumeration surveys were conducted as part of the 1950, 1960, and 1980 U.S. Censuses of Population and Housing. The design of each of these post enumeration surveys however, was considerably different. The 1950 and 1960 post enumeration surveys attempted to obtain an estimate of the total population that was "better" than the census by employing more rigorous procedures. The 1980 PES approach emphasized independence from the census rather than quality of PES enumeration.

5.21 1950 census post enumeration survey.--A sample of areas was selected for this study and intensively recanvassed. Housing units that were missed from the census were then determined along with persons residing in those units (who were by definition also missed).

A sample of housing units that were enumerated in the census also was selected and contacted in order to estimate persons missed within enumerated units. Better enumerators were used and more intensive household interviews were conducted. The PES estimated that approximately 75 percent of the gross undercount of people resulted from dwelling units having been missed. It was generally felt that these results were more a reflection of the inadequacies of the PES rather than of the true pattern of coverage error. Two general conclusions were drawn from the 1950 post enumeration survey:

(1) Despite all efforts to do a "perfect job" in the PES, people were still missed and, unfortunately, many of these same people were also missed in the census. A consequence of this is that estimates of the undercount are badly biased for certain categories, especially males, as illustrated in figure 2-5.

(2) Matching is a costly and difficult undertaking. Attempts to minimize matching costs will usually lead to substantial increases in error. It was generally felt that the net matching bias was much smaller than correlation bias and thus the overall bias was not appreciably reduced. (Generally, net matching bias is opposite in sign to correlation bias).

5.22 1960 census post enumeration survey.--The 1960 post enumeration survey consisted of two studies, each of which utilized specially trained enumerators to obtain estimates of omission and duplicate reporting of persons and housing units.

The studies consisted of:

(1) A re-enumeration of housing units in a selected sample of areas which were intensively canvassed for missed and erroneously identified housing units in the census.

(2) A sample of housing units enumerated in the census was re-enumerated to identify persons in census enumerated units who were missed in the census.

Approximately 54 percent of the persons missed were determined to have come from missed dwelling units, a result somewhat more plausible than the 1950 results. However, a comparison of the PES results with estimates from demographic analysis (see figure 2-6) indicated that post enumeration survey estimates for blacks appeared to be seriously deficient.

As in the 1950 post enumeration survey, it is generally felt that the PES technique performs satisfactorily in detecting missed dwelling units and their occupants, but does not adequately account for persons missed within enumerated units.

A major finding from this study was that estimation of the net undercount is made very difficult by missing data in the census and post enumeration survey which cause a considerable number of unresolved match status cases. Very often, this occurs in groups that exhibit the highest undercounts. If the proportion of cases for which match status cannot be determined is higher than the proportion of persons who are missed, the adjustment made for missing data will strongly affect the estimates being made.

Figure 2-5. COMPARISON OF PES ESTIMATES OF 1950 NET UNDERCOVERAGE WITH DEMOGRAPHIC ANALYSIS ESTIMATES

(population in thousands)

Source	Total	Both sexes, 0 to 14 years	Male			Female		
			15 to 39 years	40 to 64 years	65 years or older	15 to 39 years	40 to 64 years	65 years or older
Census...............	150,697	40,482	28,034	20,393	5,798	29,073	20,447	6,474
PES estimate.........	152,788	41,187	28,092	20,785	5,966	29,222	20,940	6,604
"Minimum reasonable" estimate [1].........	154,416	41,943	28,956	20,785	5,966	29,222	20,940	6,604
Coale estimate [2]......	156,130	41,970	29,340	21,110	5,960	29,610	21,540	6,600
Estimated net undercount in Census								
PES estimate.........	2,091	705	58	392	168	149	493	130
"Minimum reasonable" estimate [1]...........	3,715	1,461	922	392	168	149	493	130
Coale estimate [2]......	5,429	1,488	1,306	717	162	537	1,093	126

[1] *For these estimates, the PES estimates for children under age 10 and for males 15 to 39 were adjusted upward. Children under age 10 were adjusted upward (by .8 million) utilizing birth registration information and males 15 to 39 were adjusted upward (by .8 million) using expected sex ratios that make use of sex differentials in birth and mortality rates.*

[2] *Coale, A.J., and Zelnick, M. 1963. New Estimates of Fertility and Population in the United States. Princeton : Princeton University Press.*

*Figure 2-6. COMPARISON OF CENSUS EVALUATION STUDY (PES) ESTIMATES OF 1960 NET UNDERCOVERAGE WITH DEMOGRAPHIC ANALYSIS (DA) ESTIMATES OF 1960 NET UNDERCOUNT**

Age	White				Nonwhite			
	Male		Female		Male		Female	
	PES	DA	PES	DA	PES	DA	PES	DA
All ages.....	1.6	2.4	1.7	1.6	4.2	9.7	3.4	6.3
Under 5 years.....	1.3	1.9	1.7	1.1	2.6	6.6	1.8	5.1
5 to 9 years......	0.5	2.4	1.6	1.5	4.8	5.1	6.5	4.2
10 to 14 years....	0.7	2.5	1.7	1.5	4.3	5.0	-0.7	3.9
15 to 19 years....	1.2	3.8	1.8	2.4	-2.8	12.3	1.1	9.6
20 to 24 years....	-0.2	4.3	2.3	2.4	2.3	18.4	2.6	9.5
25 to 34 years....	1.8	3.6	1.0	1.0	2.7	18.5	4.6	6.5
35 to 44 years....	1.4	2.2	1.8	-0.2	6.4	11.5	3.4	3.8
45 to 54 years....	2.7	2.5	1.0	2.4	7.0	11.0	6.5	9.0
55 to 64 years....	2.7	0.5	2.9	1.7	6.4	8.5	4.2	11.6
65 years or older.	2.3	0.0	1.3	3.5	6.7	-5.8	2.0	2.8

**The 1960 PES estimates are for all blacks and other races but would differ trivially from those for blacks only. The PES estimates also exclude the Armed Forces Overseas but this also would have a trivial effect on the undercoverage rates.*

5.23 1980 census post enumeration survey.--The 1980 census post enumeration survey consisted of two major parts:

(1) The April and August 1980 Current Population Survey samples, which were matched to the census to obtain a match rate estimate. The Current Population Survey is a labor force survey conducted on a monthly basis.

(2) An independent sample of persons enumerated in the Current Population Survey sample was examined to determine if they were erroneously enumerated. In addition, a clerical operation was undertaken to study duplicate enumerations, and a field enumeration was made to identify errors in geographic coding.

Prior to the 1980 census PES, two pretests of the PES procedure were tested in Richmond, Virginia and Durango, Colorado. Some of the major highlights of the 1980 census pretest PES are indicated below:

(1) Richmond and Durango Pretests: Both Procedure A and B were tested. Procedure B identified approximately twice as many mover omissions as Procedure A. However, Procedure A identified approximately 3 percent more non-mover omissions than Procedure B. Thus, the overall percentage of the total omissions who were movers was about the same in the two procedures. Approximately 30 percent of the omissions in Richmond and 60 percent in Durango were movers. This illustrates the tendency for movers to be missed in the census. Overall the Procedure B omission rate was 17 percent higher than that for Procedure A in Richmond and 22 percent higher in Durango.

(2) 1980 census PES: Procedure B, Definition II was used for the 1980 census PES, in large part due to the experience in the pretests. It was generally felt that the matching problems associated with Procedure B in the pretest were manageable and this procedure was much more successful in picking up movers.

5.3 *Korea*

A post enumeration survey was used as part of the 1970 census of population. People were listed where they were staying at the time of the PES, and an inquiry was made where they were at the time of the census (Procedure B). Final results of this study were not released for general distribution; however, some general comments can be made regarding the results.

The 1970 PES census miss rate was significantly higher than PES miss rates for the censuses of 1960 and 1966. A controversy has arisen as to whether the 1970 PES estimate is "better" than the 1960 and 1966 estimates. The argument against the 1970 PES estimate revolves around the main problem with Procedure B; that is, the difficulties with matching "migrants", which has a tendency to produce erroneous non-matches, resulting in an over-estimate of the miss rates. However, it should be noted that the miss rate for nonmigrants, while less than the migrant census miss rate, was still considerably larger than the 1960 and 1966 PES estimates. The other side of the argument is critical of the methodology used in the 1960 and 1966 post enumeration surveys. These designs utilized single system estimation techniques whereby attempts were made to create PES estimates of superior quality than the census estimates (dependent systems). As indicated previously in this chapter, this type of methodology is conceptually biased as the category "missed in both the PES and the census" is not properly accounted for.

5.4 *Paraguay*

A considerable amount of PES coverage evaluation work was undertaken as part of the 1972 Census of Population. Again, the results of this work were not released for general distribution; however, considerable information regarding PES methodology may be derived from this study.

Two independent procedures were used for the PES methodology: Procedures A and B.

Comparisons of the overall census miss rates for the total population, migrants and nonmigrants, total population between types of area (urban/rural), age/sex cohorts, and household composition (head, spouse, nonrelatives, etc.) were made. In addition, estimates with and without post-stratification were made to test the effects of correlation bias in dual system estimation. Post-stratification of the estimates included the following steps:

(1) the sample data was divided into groups expected to have very different completeness rates;

(2) dual system estimates were made separately within each group and summed in order to give appropriate estimates for the nation.

Some of the major findings of the Paraguay PES were as follows:

(1) Procedure B census miss rates were higher than Procedure A estimates, although only slightly outside the confidence interval. This relationship was true for both migrants and nonmigrants when analyzed separately. Procedure A, as would be expected, picked up fewer migrants than did Procedure B; to a lesser extent it also picked up fewer nonmigrants. The latter result was likely due to the tendency of Procedure B to erroneously classify some migrants as nonmigrants.

(2) Both Procedures A and B population estimates were lower than the census counts. This was probably due to the purposive omission of certain segments of the population from the PES (for example, institutional population), as well as sampling error.

(3) The difference between incompleteness rates for Procedures A and B was primarily evident in the rural areas, where the Procedure B estimate was considerably higher than the Procedure A estimate. This could have resulted from a failure in Procedure A to pick up rural migrants, who have a tendency to be missed more frequently than urban migrants. It also could reflect a tendency for Procedure B to perform poorly where address information is of poor quality (resulting in an increased number of erroneous nonmatches).

(4) Estimates from both Procedures A and B produced census miss rates that were higher for men than women, and were roughly uniform across age categories.

(5) Miss rates in Procedure B were consistently higher than those for Procedure A for all relationship to head of household categories. This was especially true for other relatives, other nonrelatives, and employees. Since these groups are characterized by a tendency to be migrants, Procedure B may have a better chance to pick them up.

A major finding from this study was that estimation of the net undercount is made very difficult by missing data in the census and PES which cause a considerable number of unresolved match status cases. Very often, this occurs in groups that exhibit the highest undercounts. If the proportion of cases for which match status cannot be determined is higher than the proportion of persons who are missed, the adjustment made for missing data will strongly affect the estimates being made.

5.5 *Bangladesh*

Bangladesh conducted its most recent census in 1981. A post enumeration survey was conducted immediately after the census (2 weeks elapsed between the census period and the time of the PES). The PES was designed to provide information for estimating coverage errors at the national level and separately for urban and rural areas.

The sample comprised 250 enumeration areas: 150 rural and 100 urban. The sample was stratified with the primary stratification being urban/rural and the secondary stratification was accomplished by arranging the EA's according to their geographic codes. A systematic random sample of EA's was selected from this ordering and the selected EA was re-enumerated completely.

Each EA was independently matched twice by two different matchers and the results were adjudicated by a supervisor through the use of field revisits. Non-matches were verified in the field.

The principal objective was to estimate the magnitude of both the undercoverage and the overcoverage of the census. The estimate to be derived was the net coverage error rate for the census. As will be mentioned **in chapter 3,** a secondary objective of the PES was to provide indicators of the quality of the information collected in the census.

The PES was actually conducted in two stages: a PES-A field survey and a PES-B field followup operation that was also used to estimate erroneous enumerations. The data from these operations could be used to generate a PES-A, PES-B, or PES-C estimate, and the additional information on erroneous enumerations could be used to generate net coverage error rates. The findings are summarized in figure 2-7.

The census coverage estimates were significantly higher for rural areas. The completion rate of rural areas was 96.4 percent compared to 91.1 percent for urban localities. However, the erroneously enumerated rates of rural and urban areas were not significantly different. This resulted in a difference in the net error rate, with the rate being only 2.5 percent for rural areas, but 7.7 percent for urban areas. At the national level the net error rate was calculated to be 3.1 percent. This was a record accomplishment if one compares the coverage errors of all earlier censuses; they range between 8 and 15 percent.

5.6 *India*

India also conducted its census in 1981. India has a long history of conducting PES studies, with studies conducted after the 1951, 1961, and 1971 censuses. Like Bangladesh, the PES was used to measure data quality too; unlike Bangladesh, the Indian PES measured duplicates but not other types of

Figure 2-7. 1981 BANGLADESH CENSUS COVERAGE ESTIMATES AND THEIR STANDARD ERRORS

Estimate	National		Rural		Urban	
	Percent	Standard error	Percent	Standard error	Percent	Standard error
Completion rate............	95.8	0.22	96.4	0.2	91.1	1.0
Missed rate................	4.2	0.22	3.6	0.2	8.9	1.0
Erroneous enumeration rate......................	1.1	0.11	1.1	0.1	1.2	0.2
Net error rate.............	3.1	0.24	2.5	0.2	7.7	1.0

Figure 2-8. 1981 INDIA CENSUS COVERAGE ESTIMATES AND THEIR PERCENTAGE RELATIVE STANDARD ERROR

Estimate	National		Rural		Urban	
	Percent	Relative standard error	Percent	Relative standard error	Percent	Relative standard error
Net omission rate..........	18.0	4.69	15.0	6.0	27.6	7.5
Duplicates.................	2.7	*(x)*	2.5	*(x)*	3.1	*(x)*
Difference.................	15.3	-	12.5	-	24.5	-

(x) - Not available.

erroneous enumerations. Estimates were made for 16 of the 17 states in India, covering about 97 percent of the total population.

A sample of 4,000 blocks was selected, using a proportional allocation from the 16 states. Within each state the blocks were further allocated to rural and urban areas in proportion to the total number of enumeration blocks in those areas. A block averages 120 census houses. A systematic sample of blocks was drawn within each state with blocks ordered within urban/rural strata randomly.

The field interview used the PES-A concept, and an attempt was made to interview each person directly rather than collect proxy information. The findings of this study are summarized in figure 2-8.

Like Bangladesh, the coverage rate is much better for rural areas than for urban areas. But nationally and by sub-national areas, the omission rate is much higher in India, with the national rate being 18 percent. No data for previous censuses were presented.

6. REFERENCES

Bishop, Yvonne, M.M.; Fienberg, Stephen E.; and Holland, Paul W. 1975. *Discrete Multivariate Analysis*. Cambridge, Massachusetts: Massachusetts Institute of Technology Press.

Chandrasekar, C., and Deming, W.E. 1949. "On a Method of Estimating Birth and Death Rates and the Extent of Registration." *Journal of the American Statistical Association* 44:101-115.

Cowan, Charles D. 1984. *The Effects of Misclassification on Estimates from Capture-Recapture Studies*. Ph.D. diss., George Washington University.

Fellegi, Ivan P., and Sunter, Alan B. 1969. "A Theory for Record Linkage." *Journal of the American Statistical Association* 64:1183-1210.

Marks, Eli S.; Seltzer, William; and Krotki, Karol J. 1974. *Population Growth Estimation*. New York: The Population Council.

Tepping, Benjamin J. 1968. "A Model for Optimum Linkage of Records." *Journal of the American Statistical Association* 63:1321-1332.

Coale, A.J. and Zelnick, M. 1963. ***New Estimates of Fertility and Population in the United States*. Princeton: Princeton University Press.**

Chapter 3. STATISTICAL EVALUATION OF CONTENT ERROR

1. INTRODUCTION

Every phase of census data collection and processing that involves manipulation of the data has the potential for introducing errors into the census results. Besides the interviewing operation in which enumerators and respondents can make errors, there are many other operations such as editing, transcription, keying, and coding where the personnel or the procedures cause errors which affect the census content. These content errors add bias and nonsampling variance components to the total mean square error (MSE) of a census statistic. In this chapter, methods for assessing the quality of the census data and operations through the statistical estimation and analyses of these components are considered. By estimating the bias components, the magnitude of the net systematic errors which arise from one or more operations in any data collection activity can be assessed. The estimated variance components for operations are an indication of the frequency and magnitude of the variable errors which occur. Knowledge of these components, first, enables the experienced statistician to judge the quality of the operations and, in many cases, the overall quality of the collected data. Second, an analysis of these components can often point towards improvements for future censuses in the design and implementation of the operations. This results in reductions in the magnitudes of the error components and thus in better quality data. Third, the estimated components of total error can provide a more realistic measure of the accuracy of a statistic if they are appropriately combined.

1.1 *The scope of this chapter*

This chapter will emphasize two techniques which are most often used for census evaluation. These are referred to as interpenetration studies or randomized experiments and reinterview studies. The following topics will be covered:

(1) The process by which operations, respondents, training methods, collection techniques, and/or other factors contribute to the bias and variance components of the MSE.

(2) Basic techniques for modeling census error for quantitative and qualitative data.

(3) The design and implementation of studies providing data for the correlated component of response variance, simple response variance, and response bias analysis.

(4) Methods for computing and reporting the estimates of the components of MSE for census statistics.

(5) Methods of analysis of the MSE components for determining the impact of enumerator and respondent errors on the census statistics, as well as the quality of the operations and/or the responses.

The chapter begins with an introduction to modeling nonsampling error, particularly respondent errors and enumerator errors. From this theoretical basis, the effects of systematic and variable errors of census statistics can be studied, and the concepts of mean square error component analysis can be best understood.

1.2 *Sources of nonsampling error*

Nonsampling errors can be classified into two main groups: systematic (or consistent) errors and variable (or inconsistent) errors. Systematic errors are errors which occur more or less in the same direction for all the units in the sample. For example, an interviewer may consistently overestimate the value of housing for a sample of dwelling units. Systematic errors create a bias in the estimates, since their effects cumulate over the sample observations. Variable errors are errors that occur in no predictable direction and tend to approximately cancel out in fairly large samples. These are errors such as haphazard mistakes made by interviewers, keyers, and transcribers, or some careless mistakes on the part of the respondents.

Some of the main contributing sources of nonsampling error are:

(1) The frame. Part of the population may be omitted from or erroneously included in the frame; or the sampling operation, if there is one, may incorrectly sample the units.

(2) Noninterviews. Some units in the survey may not be reached during the collection period or may refuse to respond to all or part of the questionnaire.

(3) Processing operations. Keying, coding, editing, and computer programming are all potential sources of error, although quality control operations attempt to keep these errors in check.

(4) Interviewing. Interviewers may influence the respondent to give or avoid certain responses, or they may transcribe the information onto the questionnaire incorrectly. They may even make up information for reluctant respondents.

(5) Respondents. Respondents may remember certain events incorrectly or deliberately falsify some information. Or they may lack the knowledge about the information requested or misunderstand what is wanted.

(6) Questionnaire. Questions may be confusing to both the respondent and the interviewer, or they may be worded to influence the respondents to answer in a specific (and not necessarily correct) way.

(7) Weighting. The data may be improperly weighted because of processing errors or because the proper value of the weights are not known exactly.

(8) Reporting. The survey results may be misreported by the analyst or misinterpreted by the user.

In any census, evaluation is never a substitute for the control of errors. By designing quality into a census and providing quality checks at each phase, many of these errors can be avoided. Evaluation studies offer a means of determining whether the quality control programs are yielding the desired results or whether additional controls are needed. If an operation or other potential error source is suspected as being an important detriment to data quality, it can be further investigated using evaluation methods.

1.3 *The basic concepts of content evaluation*

The statistical evaluation of census content error is concerned with the estimation of the variance and bias components which, when combined, describe the total error in a census statistic. Let us first define what is meant by "the total error in a census statistic." This requires the use of a very simple model for census error.

Let N denote the number of units (that is, persons, households, housing units, etc.) in the population of interest; for example, the population may be persons in a country, city, or village. Suppose that a census is conducted and a number of questions are completed for each of the N members of the population. Consider one particular characteristic measured in the census and denote this item by c; for example, c may be income, educational attainment, age, marital status, or size of household. Now, consider the response to item c of any unit chosen from the population. Denote the unit by the index j and the response by y_j. Our simple model for describing the total error in the census is as follows:

Recorded value for unit j	=	True value for unit j	+	Error committed by respondent, interviewer, processing, etc.

which can be written more concisely as

$$y_j = \mu_j + e_j \tag{3.1}$$

where μ_j is the true value of the characteristic for unit j, and e_j is the error introduced for the unit by any number of the sources described in section 1.2. For simplicity, we shall assume each of the N units in the population responds to item c and, later, relax the assumption.

Model (3.1) forms the basis of statistical content evaluation. The model assumes that

(1) For every unit j in the population ($j = 1, \ldots, N$), a true value μ_j for the characteristic c exists.

(2) The recorded value, y_j, and the true value, μ_j, differ by an additive error term, e_j.

(3) The e_j values are random variables.

Assumption 3 means, essentially, that for each unit j there is a set of possible values that e_j may take for the census response, each having an associated probability of occurrence. For example, suppose the characteristic of interest is household gross income and consider a particular household in the census whose true income, μ_j, is 10,000. Further suppose that, for the census, a value of y_j = 9,600 was reported. Thus, the error, e_j, is 9,600 - 10,000 = -400. The model assumes that the error -400 was chosen at random from a population of possible errors for the income of household j. If the census process could be repeated for household j, there is some probability that other values of e_j would occur.

Studies conducted in the United States have shown that the errors e_j may be correlated through systematic effects introduced by census personnel. That is, the errors introduced by some operators may be large, while those introduced by other operators may be small. It is believed that the major cause of the correlation is the census enumerator. Enumerators may influence responses by rewording questions, giving positive or negative reinforcements to responses, deducing responses incorrectly, and so on.

Because of the large number of responses they collect, each enumerator can have a tremendous impact on the data. This impact can be expressed mathematically. In the appendix, models are given for describing the effect of enumerators on census and survey results. These models make the following assumption: if y_j and $y_{j'}$ are responses to an item c for two units in the same enumerator's assignment, the associated errors e_j and $e_{j'}$ are correlated. Further, the errors e_a and e_b for two responses in two different enumerators' assignments are not correlated. This essentially means that errors made by one enumerator are not affected by the errors made by another enumerator.

With the preceding assumptions, we can now discuss a measure of the total error in a census total or percent. This widely known measure is called the "mean square error."

1.4 *The mean square error of census statistics*

Let us now consider how the errors e_j affect the accuracy of census statistics. We will give the relevant formula for totals. The corresponding formulas for proportions can be readily obtained from these.

Let Y denote the observed census total for the characteristic c. Y is therefore

$$Y = \sum_{j=1}^{N} y_j \tag{3.2}$$

where y_j has the error structure given in model (3.1). The mean square error of Y, denoted by $MSE(Y)$, is defined as the sum of: (a) the square of the bias of Y, denoted by $B^2(Y)$, and (b) the variance of Y, denoted by $V(Y)$; that is,

$$MSE(Y) = B^2(Y) + V(Y)$$

By model (3.1), we have:

$$Y = \sum_{j=1}^{N} \mu_j + \sum_{j=1}^{N} e_j \tag{3.3}$$

Note that the first term, $\sum_{j=1}^{N} \mu_j$, is the true population total and is constant--that is, it has no bias or variance. Thus, the terms in $MSE(Y)$ must arise from the second term in (3.3), referred to as the response error term. In order to provide an expression for the $MSE(Y)$, we must define the expected value and variance of e_j.

In the appendix, two structures are developed for the error e_j--one appropriate for quantitative data (or data measureable on a continuous scale) and another for qualitative or categorical data. Each of these structures yields different mathematical forms for the components of the mean square error. Therefore, the interpretations of the estimated measures of bias and response variance that will be discussed later will depend upon whether the type of data being collected is continuous or categorical. A general formula for the $MSE(Y)$ is given below, which does not depend upon which of the two error structures is used. This formula, which will be discussed subsequently in detail, applies to any census total and is central to census content evaluation:

$$MSE(Y) = N^2B^2 + N(\bar{m} - 1)(CC) + N(SRV) \tag{3.4}$$

Here, B denotes the average <u>bias</u> of the y_j, CC denotes the <u>correlated component</u> of enumerator variance, $\bar{m}$ is the average size of an enumerator assignment, and SRV denotes the <u>simple response variance</u>. Equation (3.4) may be easily converted to a formula appropriate for the census mean or proportion by dividing through by N^2. The terms in (3.4) will now be defined.

1.41 <u>B, the census bias component.</u>--The bias term, B, in (3.4) is defined by

$$B = \frac{E(Y) - \sum_{j=1}^{N} \mu_j}{N} \tag{3.5}$$

where $E(Y)$ is the expected value of the total Y over the distribution of errors in Y. Conceptually, one may imagine many repetitions of the census for the same population at the same point in time. The average of the total Y over these repetitions is the expected value of Y or $E(Y)$. This bias is, therefore, a measure of the net effect of persistent, systematic errors on the total Y. From (3.5) we have that the bias of the total Y, $B(Y)$, is NB.

There are two major causes of census bias--response error and nonresponse error, which includes coverage error. For content evaluation, response error will be considered and nonresponse error will be ignored. Thus, in subsequent analyses, we shall assume that N refers to the total number of <u>responding</u> units in the population and, therefore, B is the bias resulting from <u>response</u> errors. The impact of nonresponse on census content requires information on the nonrespondents that is usually not available. Therefore, nonresponse bias cannot be estimated using the techniques to be described.

<u>Example 1.1</u>--A study was conducted to evaluate the response bias in the census classification of persons by age. A sample of 1,000 census respondents was selected completely at random, and their birth records were checked. Figure 3-1 summarizes the results for the age category "less than 14 years."

The cells of this table have been labelled (a) through (h) for use in subsequent examples.

Only cells (e), (f), (g), and (h) are needed for the current illustration.

Figure 3-1. CENSUS AND BIRTH RECORD CLASSIFICATION OF 1,000 CENSUS RESPONDENTS BY AGE

Birth Record	Census		
	Total	Less than 14 years	14 years or older
All ages....	1,000	e 290	f 710
Less than 14 years.......	g 304	a 283	b 21
14 years or older.......	h 696	c 7	d 689

By comparing the birth record classification for persons whose age is "less than 14 years" with the corresponding census classification, an estimate of B (the response biased in the census item) can be obtained. Cell (g) divided by 1,000, or .304, is an estimate of the proportion of census respondents who truly are less than 14 years of age. Cell (e) divided by 1,000, or .290, is the census-based estimate. The difference, .290 - .304 = -.014, is therefore an estimate of the bias, B. From this example, it can be concluded that the census total for persons who are less than 14 years of age is biased downward by an estimated 1.4 percentage points.

Therefore, if Y is the census total for persons whose age is 14 years or less, an estimate of $B(Y)$ is $N(.014)$. If N is 20 million persons, then it is estimated that 20,000,000 (.014) = 280,000 persons of less than 14 years of age who have been misclassified as persons of age 14 years or more.

1.42 SRV, the simple response variance.--Next, consider the simple response variance term in (3.4). *SRV* is a measure of the reliability or consistency of census responses to a specific item. It is defined as the average variance of responses to an item for the same individuals over repeated applications of the census measurement process. For example, suppose that, as part of the census, interviewers are asked to estimate by inspection the current market value of the dwellings in their assignments. For a sample of dwellings, different interviewers return to make second independent estimates of their value. Let y_{1i} denote the estimated value of dwelling i taken on the first occasion, and let y_{2i} denote the estimate taken on the second occasion. *SRV* is a measure of the differences in the estimates made on different occasions for a particular dwelling, averaged over all dwellings. It is estimated by

$$\text{(3.6)} \qquad \widehat{SRV} = \frac{1}{2}\, Avg\, (y_{1i} - y_{2i})^2$$

where *Avg* denotes the simple arithmetic mean or average over all units i in the sample.

In census evaluation studies, *SRV* is measured by reinterview studies in which interviewers revisit a sample of households from the census and reask some or all of the census questions. These reinterview responses are later compared with the original census responses, and $\widehat{SRV}$ is computed as in (3.6). An important assumption made for these studies, and one that has been much discussed in the literature (see, for example, Hansen et al. 1959), is that the observations made for the same unit on the two occasions are independent. This means that the first interview in no way influences the responses of the reinterview. When the two observations are correlated, thereby violating the independence assumption, the respondent is said to have been conditioned by the first interview. The conditioning effect in reinterview studies often leads to underestimates of *SRV*; that is the estimate of *SRV* gives a much more optimistic indication of response consistency than is actually the case. This occurs because in some cases the respondents remember the response they gave in the original interview and simply repeat it in the reinterview. Some techniques for reducing the conditioning effect in reinterview studies will be discussed in the next section.

Example 1.2--Consider the data in figure 3-1 again. Now, however, suppose that the sample of 1,000 census respondents were reinterviewed several months after the census and reasked the question on age. In this case, we want to compare the original census responses with the responses obtained during the reinterview. Therefore, replace the table heading "Birth Record" with "Reinterview" and assume that the data do not change.

As before in the reinterview study of the value of dwellings, let y_{1i} denote the response for sample person i in the census, and let y_{2i} denote the response in the reinterview. Because we are interested in the age classification of "less than 14 years," define y_{1i} as follows:

$$y_{1i} = \begin{cases} 1 & \text{if person } i \text{ is classified as less than 14 years in the census} \\ 0 & \text{if not} \end{cases}$$

and define y_{2i} similarly for the reinterview. Then, it can easily be verified that with these definitions for y_{1i} and y_{2i} the estimator of *SRV* in (3.6) can be expressed in terms of the cells in figure 3-1 as

$$\hat{SRV} = \frac{1}{2}\left(\frac{b + c}{1,000}\right) \tag{3.7}$$

In general, $\frac{b + c}{n}$, where n is the number of units in the reinterview sample, is referred to as the gross difference rate. It is the total number of discrepancies between the census and reinterview responses over the total number of persons in the sample. Conversely, $\frac{a + d}{n}$ is the rate of agreement, since cells (a) and (b) count the number of reinterview responses that are the same as the corresponding census responses.

The estimate of *SRV* from figure 3-1, assuming the second measurement is a reinterview, is

$$\hat{SRV} = \frac{1}{2}\left(\frac{21 + 7}{1,000}\right) = .014$$

Note that from the table, $\left(\frac{21 + 7}{1,000} \times 100\%\right) =$ 2.8 percent of the sample changed their responses in the reinterview.

Another measure which is often estimated from reinterview tables is the index of inconsistency denoted by I. The appendix gives the technical motivation and definition for I. An over simplified but non-technical definition of I is that I is the ratio of the *SRV* to the total variance of Y, where "total variance" includes the variability in the population of the characteristic being measured. An estimator of I is

$$\hat{I} = \frac{2 \cdot \hat{SRV}}{p_1 q_1 + p_2 q_2} \tag{3.8}$$

where p_1 and p_2 are the estimates of the proportion in the population possessing the characteristic of interest computed from the interview and reinterview data respectively and q_1 (or q_2) is $1 - p_1$ (or p_2). From figure 3-1, p_1 is $\frac{(e)}{1,000}$, q_1 is $\frac{(f)}{1,000}$, p_2 is $\frac{(g)}{1,000}$, and q_2 is $\frac{(h)}{1,000}$. The estimate of I from the table is

$$\hat{I} = \frac{2(.014)}{(.29)(.71) + (.304)(.696)}$$

$$= .067$$

A general rule for interpreting the magnitude of $\hat{I}$ is given in section 3. We shall see there that this value of I is quite small.

1.43 *CC, the correlated component of enumerator variance*.--The correlated component of enumerator variance is typically the largest and most damaging component of nonsampling variance affecting enumerator assisted censuses. Numerous studies have shown that interviewers, by the way they ask questions, probe for clarification, interpret responses, etc., can have an enormous impact on the responses elicited from respondents (see, for example, Bailar 1976, as well as Hanson and Marks 1958). Evaluation studies designed to estimate *CC* for census items can determine the extent to which enumerators are affecting the census results and thereby direct efforts to control enumerator error. The following example is useful for understanding the nature of enumerator variance.

Consider again the situation of a census in which enumerators are to estimate the value of each dwelling in their assignments. Some enumerators may tend to underestimate the values, while some may tend to overestimate them. A estimate of *CC* will tell us whether there is much difference among enumerators in these tendencies. For example, if half the enumerators underestimate and half overestimate, and if the range of error is large, then *CC* will be large. If, on the other hand, all of the enumerators err in the same direction, that is, either under- or over-estimating the values, or if there is little difference in their errors, then *CC* will be small. It is important to note that *CC* may be small even if the enumerators are contributing substantial errors to the census content. *CC* is a measure of the variability or differences in the tendencies of the enumerators to bias the responses for units in their assignments.

Now consider how to estimate *CC*. If, in the previous example, we computed the average estimated value of dwellings for the units in each enumerator's assignment and compared these, the average values might be quite different. The differences would be due, not only to enumerator biases, but also to differences in the dwellings that make up the enumerator assignments themselves. However, by implementing a special procedure for constructing enumerator assignments, a procedure called interpenetration, this latter effect which is confounding the analysis of enumerator variance can be equalized across the assignments to be compared.

Interpenetration is another word for randomization. To interpenetrate enumerator assignments simply means to randomly assign the units to be enumerated to enumerators so that, on average, each enumerator assignment is roughly balanced with respect to the socioeconomic characteristics of households and types of units. Now a comparison can be made between enumerator assignments to detect systematic enumerator errors and to estimate *CC* as well.

Estimators of *CC* look quite complex and will be covered in section 4. The following example, however, demonstrates the main ideas.

Suppose we wish to compute the correlated component of enumerator variance associated with the census classification "14 years or less." Two enumerator assignments are interpenetrated and enumerated, and the proportion of the persons in each enumerator assignment classified in the category "14 years or less" is computed. The following table summarizes the results:

Figure 3-2. PROPORTION OF PERSONS 14 YEARS OF AGE OR LESS BY ENUMERATOR ASSIGNMENT

Enumerator assignment	Number of persons	14 years or less	
		Number	Proportion
Total....	400	130	.325
Enumerator 1..	200	40	.20
Enumerator 2..	200	90	.45

The computation of *CC* from figure 3-2 is

$\hat{CC}$ = between-assignment variance - within-assignment variance

$$= \frac{\frac{1}{2}(.20 - .45)^2 - [(.2)(.8) + (.45)(.55)]}{398}$$

$$= .030$$

The first term in the computation of *CC* is a measure of the between-enumerator-assignment variance. This includes the differences between assignments resulting from the particular populations enumerated as well as from the enumerators' systematic errors. The population difference or "sampling variance" effect is measured by the second term which is subtracted leaving an estimate of the pure enumerator effect, *CC*.

A useful measure that has been used extensively in the survey evaluation literature to represent the correlation introduced by enumerators is ρ_y, the <u>intra-interviewer correlation coefficient</u> (Kish 1962). ρ_y is analogous to the measure *I*, the index of inconsistency. It is the ratio of *CC* to total variance, while *I* is the ratio of *SRV* to total variance. Just as for *I*, "total variance" is defined to include the variability in the population of the characteristic as well as nonsampling variability. An estimator of ρ_y for categorical data is

$$\hat{\rho}_y = \frac{\hat{CC}}{pq}$$

where $\hat{CC}$ is the estimator of *CC* from the interpenetrated enumerator assignments, *p* is the proportion of units possessing the characteristic among units in the combined interpenetrated enumerator assignments, and *q* is *1 - p*. For continuous data, another formula applies and is discussed in the appendix. An estimate of ρ_y from the data in this example is

$$\hat{\rho}_y = \frac{.030}{.325(1 - .325)}$$

$$= .137$$

A general rule for interpreting the magnitude of $\hat{\rho}_y$ is given in section 3. We shall see that in this example, $\hat{\rho}_y$ is enormous.

Note that a reinterview or any second measurement per unit is not required to compute *CC*. This makes enumerator variance studies less expensive than studies to estimate the bias, *B*, or the simple response variance, *SRV*. Some of this cost savings is lost to interpenetration, however, as a result of increased travel for enumerators over the area interpenetrated. In the next section, we will consider interpenetration study designs which attempt to minimize travel costs.

1.44 <u>*MSE(Y)*: putting it all together</u>.--Now that we have explored the meanings of the components of the mean square error of the census total given in (3.4), let us see how these components--*B*, *CC*, and *SRV*--combine to determine the total error of *Y*.

It is interesting to consider which of the three components is the most important, or rather, which one has the greatest potential for substantially increasing *MSE(Y)*. As we see from formula (3.4), the precise answer to this question depends upon several factors: besides the <u>types</u> of errors committed, it depends upon (a) the size of the population being investigated (or equivalently, the size of the area for which census statistics are being reported), (b) the average size of the enumerator assignments, and (c) the particular characteristic being reported (that is, the size of *Y*).

Of all the components, bias is perhaps the easiest to discuss since we can readily imagine the impact of a 3 percent or 5 percent bias on the census results. Therefore let us consider, in terms of its impact on the *MSE(Y)*, how much enumerator variance is equivalent to some level of bias. Similarly, we will consider how much simple response variance is required to equal the impact on *MSE(Y)* of a given level of bias. Figures 3-3 and 3-4 give the equivalent levels of enumerator variance (as measured by ρ_y) and simple response variance (as measured by *I*) for levels of <u>relative bias</u> ranging from 1 percent to 5 percent.

In figure 3-3, the population size is 1 million persons (perhaps a moderate size city), and in figure 3-4, the population size is 5,000 persons (a small village). In both cases, the characteristic being measured is of the categorical or qualitative type with two categories in which one half of the population possesses the characteristic.

Figure 3-3. EQUIVALENT EFFECTS OF ENUMERATOR VARIANCE AND SIMPLE RESPONSE VARIANCE, BY LEVEL OF BIAS, FOR LARGER POPULATIONS

(N=1 million)

$B(Y)$ (as a percent of Y)	Equivalent ρ_y			Equivalent I
	$\bar{m}$=250	$\bar{m}$=500	$\bar{m}$=1,000	
1 percent..	.402	.200	.100	none
2 percent..	none	.800	.400	none
3 percent..	none	none	.900	none
4 percent..	none	none	none	none
5 percent..	none	none	none	none

Figure 3-3 shows the importance of bias for large populations. For example, to equal "the impact of a 1 percent bias, a ρ_y of at least .4 is required when enumerator assignments average 250 persons and at least .1 when $\bar{m}$ is 1,000." For populations of size 1 million, it is impossible for simple response variance to be as serious as a bias of 1 percent or larger. Further, it is impossible for enumerator variance to equal the impact on *MSE(Y)* of a 4 percent bias unless the average size of enumerator assignments is very much larger than 1,000 persons. It is important to note that the impact of ρ_y increases as $\bar{m}$, the average assignment size, increases. This can be readily seen in formula (3.4) and in figures 3-3 and 3-4.

Figure 3-4 might apply for a small reporting area such as a village or some small subgroup of a larger population. Now bias may no longer be the most critical component as for the larger populations. Enumerator variance can often achieve the same level of impact as a 5 percent response bias. For the sizes of $\bar{m}$ and ρ_y often reported in the literature, values of .001 to .05 are not uncommon. The index of inconsistency, however, even for a small population, can seldom achieve the impact of *MSE(Y)* of small values of relative bias.

Figure 3-4. EQUIVALENT EFFECTS OF ENUMERATOR VARIANCE AND SIMPLE RESPONSE VARIANCE, BY LEVEL OF BIAS, FOR SMALLER POPULATIONS

(N=5,000)

$B(Y)$ (as a percent of Y)	Equivalent ρ_y			Equivalent I
	$\bar{m}$=250	$\bar{m}$=500	$\bar{m}$=1,000	
1 percent..	.002	.001	.0005	.5
2 percent..	.008	.004	.002	none
3 percent..	.018	.009	.005	none
4 percent..	.032	.016	.008	none
5 percent..	.050	.025	.013	none

There is some recent evidence in the literature that large values of ρ_y, and B are typically found when I is large. This means that a large simple response variance may be an indicator of large enumerator variance and/or large response biases. Thus, estimating I for a census may yield some information of the two large components of *MSE(Y)*. More work is needed to test this conjecture.

To summarize, we have illustrated the relative effects of *B*, *CC*, and *SRV* on *MSE(Y)* for censuses. For large populations, response bias is the most important component and *CC* and *SRV* may be unimportant. For small populations, the impact of *CC* can be considerable and may often be the largest component of mean square error. The impact of *SRV* increases as the size of the population decreases, but is usually less important than *B* or *CC*. However, there is some evidence that a large *SRV* may be an indicator of large enumerator or response bias effects. In addition, *SRV* may be the most convenient parameter to estimate.

2. DESIGN OF STUDIES TO ESTIMATE THE COMPONENTS OF CENSUS ERROR

This section presents some general principles for good evaluation study design. First, it deals with enumerator assignment "interpenetration" studies for estimating the correlated component of enumerator variance, *CC*. Interpenetration is the technique of combining together two or more enumerator assignments and then reassigning the units to the same enumerators using some method of randomization. In this way, each new assignment is said to be "statistically identical." That is, on the average, the procedure produces assignments which have the same distributions of population characteristics. The enumerator assignments which have been interpenetrated can be compared with respect to the average of characteristic units in order to detect enumerator differences.

Interpenetration studies may be expensive and difficult to control. Since the interpenetrated enumerators' assignments are spread over a much larger geographic area, enumerator travel is increased. In addition, special records must be kept so that the enumerator(s) associated with every unit in the interpenetrated assignments can be identified.

Depending on their design, reinterview surveys may be used to estimate either the simple response variance, *SRV*, or the response bias, *B*. The reinterview survey designed to estimate *SRV* is basically a repetition of the census for a sample of units. In our model of the last section, it was assumed that the second interview is an <u>independent</u> replication of the first interview. In practice, this is very difficult to achieve for reasons discussed in section 2.2.

The reinterview survey designed to estimate *B* is considerably different than the former type. For this survey, we aim to obtain the "truth" in the reinterview. This calls for innovative methods to help the respondents better recall events or to enhance their understanding of the questions. These methods may be infeasible for use in the full census because of their cost and/or complexity. They are considered for reinterviews because of their potential for improving responses. In addition, better enumerators and improved field procedures are used to decrease the errors occuring in field operations. *SRV* may also be estimated from this type of reinterview, although the estimator is more complex than the estimator from a replication-type reinterview. (See section 4.2 in the appendix.)

It is not uncommon for reinterview surveys to incorporate both objectives--to estimate *SRV* for some items and to estimate *B* for other items. In these surveys, the items for which *SRV* is wanted are reasked exactly as worded in the census, and no improvements in the field procedures are attempted. For the items for which *B* is to be estimated, probing questions, better training, and other enhancements to the interview and field procedures may be used in order to get the best response possible.

2.1 *Interpenetration of enumerator assignments to estimate the correlated component*

For any evaluation study to produce meaningful results, it must be properly planned and implemented. It is of primary importance that the statistical concepts be thoroughly understood and appreciated, and that the conditions under which the study must be carried out be fully supported by people at various levels in the statistical office. Knowledge of the assumptions made in the model development of the estimators, and of the consequences of deviating from these assumptions in the conduct of the study, is an important requirement for the study planners. Many times during an interpenetration study, the project leaders must make decisions which affect the randomization of assignments, the independence of the enumerator assignments, or the environment of the study personnel. Uninformed decisions threaten the validity of the correlated component estimates. In these situations, the guiding principle should be to meet the objectives of the study without incurring excessive costs.

A major danger in attempting to measure the error in a survey by interpenetration studies is the presence of a "study effect." This effect, which is unavoidable, is due to changes in the procedures or even in the attitudes of the enumerators as a result of the presence of the evaluation study. If these effects are large, it could render the evaluation results useless, since the measured effects are not indicative of the total census enumeration effects but only of the study enumerator effects. For this reason, it is important to maintain conditions in the study sites which are as far as possible identical to conditions in field offices not in the study.

At the same time, the conditions assumed for the model must also be met. The enumerator assignments, in order to be randomized or interpenetrated, may require more work on the part of the enumerators. For example, interpenetration may require more travel for the study enumerators than is typical for the enumerators not in the study. The effects of this on the results must be monitored and controlled.

Other problems affect the independence of the enumerator assignments. For example, an enumerator in the study is unable to complete his/her assignment; assigning another <u>study</u> enumerator to the assignment will introduce a correlation between the two assignments which is in violation of the model assumptions. Or, if refusal cases for all enumerators are handled by a small crew of "refusal conversion specialists", a correlation is introduced between assignments. These cases must then be excluded in the analysis. However, if refusal rates are high, this also creates problems in the analysis.

There are no perfect solutions to these problems; however, a knowledgeable statistician can usually develop solutions which are acceptable operationally and come as close as possible to the theoretical ideal. Now let us consider the stages involved in the implementation of a study of enumerator correlated error.

2.11 <u>Planning</u>.--Preliminary to the design of an evaluation study, five steps should be performed. These are:

(1) <u>Objectives of the evaluation</u>. Develop a specific and very descriptive statement of the objectives. Often, during planning, it is easy to become engrossed in the details and forget the main objectives of the study. This is often the cause of poor decisionmaking.

(2) <u>Method of evaluation</u>. If correlated or systematic enumerator errors are the concern, then an interpenetrated design for estimating enumerator error is appropriate. However, for some characteristics, sucn as sex, one might expect the errors to be more random in nature and systematic differences between enumerators to be unimportant. For these characteristics, estimating the correlated component may yield very little about the importance of errors in the census, and an interpenetration study is not appropriate. Thus, determine whether estimating the correlated component is important for the characteristics being measured before deciding to perform an interpenetration study.

(3) <u>Data to be collected</u>. Determine what data are relevant to the purposes of the evaluation. For example, data must be collected to allow the survey data to eventually be linked to a specific enumerator; records must be kept on units which should be deleted from the study because procedures were not followed for them; other data may be necessary in order to aid in the interpenetration of estimates. These should all be prespecified.

(4) <u>Precision desired</u>. The specification of the degree of precision wanted in the results is an important step. Depend upon previous studies to determine the approximate size of ρ_y to be estimated and specify the desired coefficient of variation for the estimate. Occasionally, it may be decided that the costs involved are too great to conduct the study as originally planned.

(5) <u>Study design</u>. Specify the number of collection offices to be involved in the study, the number of enumerators and the

way their assignments will be controlled, the interpenetration scheme and how to handle special problems, and the timing of the study.

Is the design feasible? For example, interpenetration of enumerator assignments often means more travel for the enumerators. In fact it can be shown that enumerator travel increases in proportion to the square root of the number of enumerators being interpenetrated in an area. (When pairs of enumerators' assignments are interpenetrated, each interviewer travels approximately 2 = 1.41 times as far.) In some areas, this may not be feasible considering the timing of the survey.

The steps involved in the sample design for an enumerator variance study are now listed.

(1) Using the formulas given in the appendix for the precision of ρ_y, determine the required number of enumerator assignment pairs to be interpenetrated (this is ℓ in formula (4.1.15)). The maximum acceptable relative variance is .25. A larger relative variance would not allow the detection of enumerator effects.

(2) Determine the number of collection offices (study sites) to be involved in the evaluation, using the following as a rough guide:

$$\text{number of study sites} = \ell \div 40$$

For example, if the number of enumerator pairs from step (1) is 500, then 500/40 or 13 study sites are needed. This number can be increased/decreased depending upon the particular characteristics of the census collection offices. However, a number smaller than 10 study sites usually is not recommended.

(3) Prepare a list of the census collection offices ordered by urban and rural characteristics and, within these classes, ordered by size. In addition, other variables that are believed to be correlated with enumerator error may be used to sort the offices within size classes.

(4) Select a systematic sample of collection offices from the sorted list. These will be referred to as study sites. If the collection offices vary considerably in size, a more sophisticated "unequal probability" sampling scheme might be more efficient.

(5) For each study site, prepare a list of enumeration assignment pairs by pairing together assignment areas which are geographically adjacent to one another.

(6) For each study site select a sample of ℓ/ (number of study sites) enumerator assignment pairs. These assignments will be referred to as the study assignments.

(7) For each pair of study assignments, prepare a list of the dwelling units located within the assignments and sort the list geographically.

(8) Systematically assign the first, third, fifth, etc., unit to enumerator A of the pair, and the remaining units to enumerator B of the pair. This completes the interpenetration design.

2.12 Preparation.--Before the data are collected, the general nature of the results and the way in which they will be analyzed should be described. This includes identification of charts, graphs, and tables to be constructed and the timing of the reports.

Procedures for clerks, quality control personnel, and supervisors should be written, and the appropriate training materials should be developed. In addition, it is important that these procedures be pretested. This nearly always results in improvements in the procedures or clerical forms. It can often reveal that the cost will be much greater than expected.

Generally speaking, the enumerators to be evaluated should receive the same training as the other enumerators in order to minimize the study effect. However, often it is necessary to provide additional training to these enumerators to cover changes in the handling of their assignments. Care must be taken to minimize the administrative differences between study enumerators and other enumerators.

During training of the study support personnel, stress the importance of following procedures precisely, even if they seem inefficient. Since the objectives of interpenetration are not readily understood by clerical personnel, they may adopt procedures which are contrary to these objectives.

2.13 Implementation.--During the conduct of the study, it is useful for the study planners and project leaders to visit the work sites, both in and out of the study, to insure that the proper conditions are being maintained in the study sites.

2.14 Analysis.--During the documentation of the results, each stage of the experiment should be fully described. Be sure to document the specific problems encountered in the execution and data processing phases. Include estimates of the standard errors of the estimated, ρ_y's. (See section 3.)

2.2 *The design and implementation of reinterview surveys*

The planning required for a large scale reinterview or post-enumeration survey can be as complex as that required for a large scale sample survey. In fact, because of certain distinctive features peculiar to its purpose and content, the evaluation survey requires the balancing of many aspects of cost and accuracy normally not encountered in sample surveys, as this section will demonstrate.

The two alternative objectives for a reinterview survey are: simple response variance objective--which strives for an independent repetition of the survey or census under the same general conditons--and response bias objective--which strives for a measure of truth or the most accurate response obtainable from a respondent. The cost, planning, and implementation of a reinterview survey for each objective is quite different as might be expected. The first objective usually requires the same questions about characteristics, the same method of obtaining and recording responses, and the same sponsorship of the survey as for the census. By contrast, the second objective requires results to serve as a standard of measurement. Here the deficiencies of the census are minimized in the reinterview survey by application of more rigorous field procedures such as using better trained and more highly qualified interviewers, choosing the most knowledgeable respondent to provide the data, applying detailed questioning sequences to probe areas where questions or instructions have been ambiguous or inadequate, and reconciling different responses collected in the two interviews. The latter type of reinterview deliberately changes the questions and techniques in an attempt to shed light on errors arising from problems in questionnaire wording, enumerator and/or respondent failure, etc. Despite the distinctions between the two objectives, there are some commonalities. These common aspects are discussed first, followed by the special requirements and considerations for each objective.

2.21 Sample design.--The sample design for reinterview surveys is usually similar to that of large sample surveys. The design may be a complex sample design which is intended to minimize the cost of travel for the reinterview while still satisfying the precision requirements. The sample size is determined by the precision requirements for the estimates of response bias and simple response variance (see the appendix).

Some reinterview surveys of the U.S. Census Bureau serve a dual role. As a check on interviewer performance, they are a means of interviewer quality control. In addition, they provide data for estimation of the response variance component. For censuses, the reinterview sample designs are usually multi-stage stratified designs which parallel those of ongoing current surveys. In addition, current survey interviewers may be called upon to collect the reinterview survey data. The reinterview sample design often specifies that households within a primary sampling unit be selected with probability inversely proportional to the number of households in the primary unit. Since primaries are typically selected with probabilities proportional to size, this makes the reinterview sample design self-weighting. Since the error of respondents is of primary interest, only respondents or responding households in the census are eligible for the reinterview sample, and non-respondents are excluded. The theory of the previous sections would then be applied to the reinterview subsample.

2.22 Matching.--Once the sample has been drawn and the reinterview conducted, the reinterview households must be matched to the survey households so that the case-by-case comparisons can be made. Since only the matched sample cases are used, it is important that the matching criteria be such as not to bias the matched comparisons. This means the proportion of erroneous matches should be kept at a minimum, since for an erroneous match the survey and reinterview characteristics will tend to differ more than for a true match--that is, erroneous matches tend to bias the simple response variance or response bias measures upward. However, there is also a danger in making the matching criteria too strict, since the opposite effect could occur. That is, those cases which can be matched by the strict rule may show relatively small differences between the reinterview and the survey. Thus the nonsampling error measures are biased downward.

2.23 Timing.--How soon after the original interview should the reinterview be conducted? If a reinterview is conducted soon after a case has been interviewed, the effect of conditioning is worsened. This conditioning affects the between interview correlation, so that estimates of response variance are biased. It may also affect the reinterview survey's ability to obtain the true response, if that is the objective. There is also a danger that a household will be reinterviewed before the census interview has been closed out, unless the reinterview is delayed at least until all census interviews have been completed. On the other hand, a late start also has the potential for loss of accuracy. Respondents would be questioned about events increasingly remote in time, and there may be a problem with recall loss. Or, a bias mav affect the estimates as a result of some sample respondents moving before they can be reinterviewed.

Studies have shown that, for most characteristics, delays of up to 3 months have no identifiable effect on the data. However, for items subject to recall loss--income, mobility, victimization--a reinterview closer in time to the original interview produces better results. (See [Bailar 1968] for further discussion of this topic.) For the U.S. Census Bureau, the timing of the reinterview may range from 1 week (for an ongoing survey) to several months (for a census) from the original interview.

2.3 *Special considerations for reinterview surveys with the simple response variance objective*

If the objective is to measure simple response variance, ideally the reinterview survey should be an identical repetition of the census. Of course, this is not possible because of:

(1) Differences in Scope. The reinterview survey is much smaller than the census, which affects the administration and interviewer workloads.

(2) Differences in Purpose. There is usually less importance attached to the reinterview survey, which is used only for evaluation purposes, than to the census.

(3) Conditioning. The responding households have been affected by the original interviews, and this could affect responses.

(4) Timing. The reinterview survey occurs at a time later than the original survey, and this could affect responses.

The utility of the estimates of simple response variance depends upon the success of the survey designers in minimizing the impact of these differences. This requires an uncompromising attitude toward maintaining the same standards and conditions operating in the census.

2.31 Questionnaire.--Ideally, the reinterview questionnaire should be a copy of the census questionnaire except for minor changes to the introduction. Also, in order for the reference periods to be the same in both the census and the survey, some changes may be required due to the later interviewing period. However, in the past, other liberties have sometimes been taken, such as shortening the interview length to save on respondent burden and to focus the study on a small subset of characteristics.

2.32 Interviewers.--The selection, qualifications, training, and supervision of the interviewers should be as close as possible to those in the census. To avoid further dependence between the census and the reinterview, the interviewer who obtained information from a household in the census should not be allowed to reinterview the same household.

2.33 Respondents.--The respondent rules used in the census also apply here. Sometimes, however, to minimize the conditioning effect of the census, the original respondent is often accepted only as a last resort. Whether this is feasible for any particular reinterview survey depends upon the type of data that are being collected. For example, there may be concern about the accuracy of proxy information which would advise against the practice of avoiding the original respondent.

2.4 *Special consideration for reinterview surveys with the response bias objective*

Although its objective is to provide a measure of the "true" characteristic of every individual in the subsample, the reinterview, at best, provides only "better" responses than those obtained in the original interview. This means the interpretation of the results should follow the theory given in the appendix for Case 2 in section 4.23, which yields rough approximations to the measures identified for Case 3.

The chief disadvantage of this method of evaluation is that the study analyst never knows how closely the response bias objective (Case 3 in the appendix) was approximated in the reinterview. A respondent who reports an incorrect age, income, etc., in the census also tends to do so in the reinterview. Further, when there are differences between the census and the reinterview, it is usually not possible to determine which value is "better."

Another disadvantage of the method is cost; however, this may be more a disadvantage for surveys rather than censuses. Much effort and money are devoted to better training and interviewing, reconciliation of discrepancies, and better quality control during reinterview data collection and processing. Since most surveys have a fixed overall expenditure limit, increased expense to measure the accuracy will usually require curtailed expenditures in the main survey with consequential decreased accuracy--for example the overall sample size may have to be decreased. Thus, some balance is required between the objectives of measurement of the population characteristics and measurement of the accuracy of the measurements.

In summary, the usefulness of the reinterview results depends upon the care taken to create the ideal survey conditions for accurate measurement. Only then can users of the results be confident that the estimated measures of bias and variance are reasonable approximations to the real levels of survey error.

2.41 Questionnaire.--In designing the reinterview questionnaire to estimate bias, emphasis must be placed on investigating each variable thoroughly. This often means restricting the number of variables to be checked in order not to have an interview of unreasonable length. For example, in the U.S. census, the question in figure 3-5 appeared for a household person in the regular census. For the reinterview, the question in figure 3-5 was replaced by the sequence of questions shown in figure 3-6.

Figure 3-5. 1980 U.S. CENSUS QUESTIONNAIRE

14. What is this person's ancestry? *If uncertain about how to report ancestry, see instruction guide.*

(For example: Afro-Amer., English, French, German, Honduran, Hungarian, Irish, Italian, Jamaican, Korean, Lebanese, Mexican, Nigerian, Polish, Ukrainian, Venezuelan, etc.)

2.42 Reconciliation.--As mentioned above, when discrepancies between the census responses and the reinterview response are found, it is usually not possible to assume that the reinterview response is the correct one. Thus, in reinterview surveys with the response bias objective, the approach has been to try to reconcile discrepancies by determining which of the two entries is "true", or, if neither is true, what is "true." For example, the interviewer might ask the respondent: "In the census, you indicated that your ancestry is American but your reinterview responses

Figure 3-6. 1980 U.S. CENSUS CONTENT REINTERVIEW QUESTIONNAIRE

Many people have ancestors that were born in other countries. We would like to ask you about the countries in which your ancestors were born.

19. In what country were the following ancestors on your father's side born?		**20. In what country were the following ancestors on your mother's side born?**		**21. CHECK ITEM B** *Refer to items 19 and 20.*
ANCESTOR	COUNTRY	ANCESTOR	COUNTRY	
Your father		Your mother		1 ☐ ALL responses to 19 and 20 indicate "U.S.A." or "Don't know" – *Continue with 22 on page 8*
Father's father		Mother's father		
Father's mother		Mother's mother		2 ☐ At least one response indicates a country other than U.S.A. – *SKIP to 25 on page 9*
Earlier generations on your father's side		Earlier generations on your mother's side		

22. All of the ancestors you have told me about have been American (or don't know). In what country were your ancestors who first came to the United States born?	**23. CHECK ITEM C** *Refer to item 22*	**24. Which one of the countries you reported best describes your ancestry?** *If necessary, read all responses to item 22.*
____________ ____________ 0 ☐ Don't know	1 ☐ Single, specific country – *SKIP to 25 on page 9* 2 ☐ More than one country – *Continue with 24* 0 ☐ Don't know – *SKIP to 25 on page 9*	____________ 0 ☐ Don't know

indicate that you are of Polish ancestry. Could you tell me which is correct?" Of course, there is no proof that this method yields correct responses. In fact, it is suspected that respondents may tend to support their most recent responses to avoid embarrassment.

Another problem which has considerable ramifications is the question of whether or not the reinterviewer should be provided with the results of the census. This would be desirable, since it would allow the reinterviewer to check his/her own answers against the original entry and to determine more positively what the correct answer is. The potential disadvantage is a tendency of the reinterviewer to simply confirm the original entry even when it is wrong, introducing a large positive covariance between the trials. The alternative is a more costly independent reconciliation procedure. Here the second interviewer, the reinterviewer, does not have access to the original results during the reinterview. The census and reinterview results are later compared by office personnel who identify any differences. Later a third interviewer conducts a reconciliation interview at the households with discrepant responses.

Research has shown that, for many items, the dependent reconciliation approach yields the same results as the independent approach. However, for items such as school enrollment, educational attainment, and income items, dependent reconciliation could have an effect on measures of the response bias. The general consensus among survey analysts is that the cost of an independent reconciliation procedure is usually not justified for reinterview surveys. (See [Bailar 1968] for further discussion.)

2.43 Interviewers.--Because of the nature of the reinterview, only interviewers of the highest competence should be reinterviewers. One difficulty is simply defining the criteria to be used to identify such interviewers. It is generally agreed that interviewing experience is an essential requirement, particularly experience with concepts to be covered in the reinterview. Other qualities thought to be desirable for interviewers are intelligence, alertness, ability to get people to talk freely, etc. Perhaps the best pool of candidates is the group of enumerators in the census. The advantages are that time and money could be saved in training and that there will usually have been some opportunity to observe them in an actual interviewing situation.

2.44 Respondents.--The most knowledgeable respondent about personal characteristics is usually the person itself (except for minors, mentally incompetent individuals, and similar cases). However, again, cost is the limiting factor. In one study at the U.S. Census Bureau, this so-called self response procedure resulted in an increase in callbacks of about 50 percent (Marks et al. 1953). There is also evidence that strict adherence to this procedure may lead to decreased response rates.

Often a compromise procedure is used. Instead of requiring that the respondents respond for themselves, there is an order of preference:

(1) The person for whom the information is being obtained,

(2) The original respondent,

(3) Another respondent meeting acceptability standards.

2.45 Processing.--Care should be taken at the processing stage not to introduce coding or keying errors into the results, since even a relatively small level of error may have a considerable effect on the measures of bias. In most cases, the usual quality control procedures used for the main survey data are not adequate, and more stringent checks are required.

3. *TABULATING, REPORTING, AND INTERPRETING THE RESULTS*

This section describes the mechanics of estimating from evaluation studies the response error measures presented in the previous sections. The computational formulas are given for estimating: CC and ρ_y from enumerator assignment interpenetration studies, B from reinterview studies with the response bias objective, and SRV from simple response variance reinterview studies. Also included are computational formulas for estimating the standard errors of these estimators. A recommended structure for reporting the results of evaluation studies is also discussed with illustrations of the standard formats for presenting summary tables of the results. The final section treats the interpenetration of the evaluation findings and presents some simple rules of thumb for gauging the magnitudes of $\hat{\rho}_y$ and $\hat{I}$.

3.1 *Estimation and tabulation*

Throughout this section, we will be presenting computational formulas for the estimation of the various response error measures described in section 1. The derivations of the analytical forms of the estimators and the appropriate model assumptions are given in the appendix. Some familiarity with the results of the appendix would be useful but is not absolutely required for this section.

None of the estimators dealt with here will require the use of sample weights. This greatly simplifies the presentation of the formulas and the computations of the estimate.

3.11 The estimation of CC and ρ_y in enumerator interpenetration studies.--The following additional notation will be required in this section:

Let

ℓ = the number of pairs of interpenetrated enumerator assignments in the study

A,B = the two enumerators associated with an interpenetrated assignment pair; enumerator A is assigned one (random) half of the assignment pair and enumerator B is assigned the remaining half

$Y_{Aj}(h)$ = the unweighted total number of elements classified in category j for enumerator assignment A of pair h, for $h=1, \ldots, \ell$

$Y_{Bj}(h)$ = the corresponding total for enumerator B's interpenetrated assignment for pair h

$m_A(h)$ = the total number of elements classified in enumerator A's interpenetrated assignment in pair h

$m_B(h)$ = the corresponding total for enumertor B in pair h

Y_j = the total for category j over the n elements in all interpenetrated assignments for the study

Figure 3-7 is a classification table that summarizes the totals needed for the computation of $\hat{CC}$ and the estimator of its variance. C is the number of categories for the census item to be evaluated. For example, for the item sex, C = 3: male, female, or unreported; for the item marital status, C is 5: single, married, divorced/separated, widowed, or unreported. For each interpenetrated assignment pair h $(h=1, \ldots, \ell)$, totals for each of the C categories are computed over the $m_A(h)$ elements (persons, housing units, etc.) in enumerator A's assignment and

Figure 3-7. NOTATION FOR COMPUTING THE CORRELATED COMPONENT ESTIMATOR AND ITS STANDARD ERROR (General procedure for a given pair of interpenetrated assignments)

Interpenetrated Enumerator Assignment	Total Elements	Census Classification ($j = 1, \ldots, c$)			
		Category 1	Category j	(Etc.)	Category c
Total.......	n	Y_1	Y_j	(Etc.)	Y_c
Pair 1:					
Enumerator A........	$m_A(1)$	$Y_{A1}(1)$	$Y_{Aj}(1)$	(Etc.)	$Y_{Ac}(1)$
Enumerator B........	$m_B(1)$	$Y_{B1}(1)$	$Y_{Bj}(1)$	(Etc.)	$Y_{Bc}(1)$
Pair h:					
Enumerator A........	$m_A(h)$	$Y_{A1}(h)$	$Y_{Aj}(h)$	(Etc.)	$Y_{Ac}(h)$
Enumerator B........	$m_B(h)$	$Y_{B1}(h)$	$Y_{Bj}(h)$	(Etc.)	$Y_{Bc}(h)$
Pair ℓ:					
Enumerator A........	$m_A(\ell)$	$Y_{A1}(\ell)$	$Y_{Aj}(\ell)$	(Etc.)	$Y_{Ac}(\ell)$
Enumerator B........	$m_B(\ell)$	$Y_{B1}(\ell)$	$Y_{Bj}(\ell)$	(Etc.)	$Y_{Bc}(\ell)$

likewise for enumerator B of the pair. Figure 3-8, then, gives the formulas for the computation of $\hat{CC}_j$, the estimator of CC, category j and ρ_{yj}. Figure 3-9 gives the corresponding formulas for 95 percent confidence intervals on the estimates. For each formula given, an illustration of its use based on the data in figure 3-10 is provided.

3.12 Estimation of *B* and *I* in reinterview studies.--Now consider the formulas for computing the response error measures appropriate for reinterview studies. Figure 3-11 sets forth the notation needed to describe the computations. The cell entries are unweighted totals of the observations on the sample elements selected for the reinterview study. For this table:

n = the total number of reinterview cases for which there was a report in both the census and the reinterview

c = the number of categories for the characteristic

Y_{ij} = the number of (unweighted) sample elements in the i^{th} category in the reinterview and j^{th} category in the census

$Y_{.j}$ = the total number of sample elements in the j^{th} category in the census

$Y_{i.}$ = the total number of sample elements in the i^{th} category in the reinterview

Figure 3-12 gives the data used in the example computations. Figure 3-13 shows the formulas and computations for estimating the

Figure 3-8. COMPUTING $\hat{C}C$ AND $\hat{\rho}_y$

<u>Correlated component for category j:</u>

$$\hat{C}C_j = \frac{1}{\ell}\sum_{h=1}^{\ell} \hat{C}C_j(h)$$

where

$$\hat{C}C_j(h) = \left(\frac{n_h-1}{n_h-2}\left[\frac{Y_{Aj}^2(h)}{m_{Ah}} + \frac{Y_{Bj}^2(h)}{m_{Bh}}\right] - \left[Y_{Aj}(h) + Y_{Bj}(h)\right] \times \left[\frac{Y_{Aj}(h) + Y_{Aj}(h)}{n_h} + \frac{1}{n_h-2}\right]\right) \Big/ m_h$$

for $\quad n_h = m_{Ah} + m_{Bh}$ and $m_h = n_h - \dfrac{(m_{Ah}^2 + m_{Bh}^2)}{n_h}$

<u>Correlated component for married category (subscript $j=1$ assumed):</u>

$$\hat{C}C = \frac{1}{3}\,(.0003 + .008 + .021) = .01$$

where

$$\hat{C}C(1) = \left(\frac{942}{941}\left[\frac{(312)^2}{480} + \frac{(290)^2}{463}\right] - \left[312 + 290\right] \times \left[\frac{312 + 290}{943} + \frac{1}{941}\right]\right) \Big/ 471$$

$$= .0003$$

$$\hat{C}C(2) = \left(\frac{1,630}{1,629}\left[\frac{(530)^2}{841} + \frac{(395)^2}{790}\right] - \left[530 + 395\right] \times \left[\frac{530 + 395}{1,631} + \frac{1}{1,629}\right]\right) \Big/ 815$$

$$= .008$$

$$\hat{C}C(3) = .021$$

<u>Intra-enumerator correlation coefficient for category j:</u>

$$\rho_{yj} = CC_j \Big/ \left[\frac{Y_j}{n}\left(1 - \frac{Y_j}{n}\right)\right]$$

<u>Intra-enumerator correlation coefficient for married category:</u>

$$\rho_y = \frac{.01}{\frac{2,430}{4,043}\left(1 - \frac{2,430}{4,043}\right)} = .04$$

Figure 3-9. COMPUTING THE STANDARD ERRORS AND NINETY-FIVE PERCENT CONFIDENCE INTERVALS

Ninety-five percent confidence limits for the correlated component for category j are:

$$\hat{CC}_j \pm 2\sqrt{\frac{1}{\ell(\ell-1)} \sum_{h=1}^{\ell} (\hat{CC}_j(h) - \hat{CC}_j)}$$

Ninety-five percent confidence limits for the correlated component, married category are:

$$.01 \pm 2\sqrt{\frac{1}{3(2)} \left[(.003 - .01)^2 + (.008 - .01)^2 + (.021 - .01)^2\right]}$$

$$.01 \pm .012$$

Ninety-five percent confidence limits for the intra-enumerator correlation coefficient for category j are:

$$\hat{\rho}_{jy} \pm 2\sqrt{\frac{1}{\ell(\ell-1)} \sum_{h=1}^{\ell} \left(\hat{\rho}_{yj}(h) - \hat{\rho}_{yj}\right)^2}$$

where

$$\hat{\rho}_{yj}(h) = \frac{\hat{CC}_j(h)}{\left[\frac{Y_{Aj}(h) + Y_{Bj}(h)}{n_h}\right]\left[1 - \frac{Y_{Aj}(h) + Y_{Bj}(h)}{n_h}\right]}$$

Ninety-five percent confidence limits for the intra-enumerator correlation coefficient, married category:

$$.04 \pm 2\sqrt{\frac{1}{3(2)} \quad (.001 - .04)^2 + (.03 - .04)^2 + (.08 - .04)^2}$$

$$.04 \pm .046$$

where

$$\rho_y(1) = \frac{.003}{\left[\frac{312 + 290}{943}\right]\left[1 - \frac{312 + 290}{943}\right]} = .001$$

$\rho_y(2) = .03$ (computed as for $\rho_y(1)$)

$\rho_y(3) = .08$ (computed as for $\rho_y(1)$)

Figure 3-10. EXAMPLE OF PROCEDURE FOR THREE INTERPENETRATED ENUMERATOR PAIRS (Table contains artificial data for marital status)

Interpenetrated Enumerator Assignment	Total Elements	Census Classification ($j = 1, \ldots, 5$)				
		Married	Widowed	Divorced or Separated	Never Married	Not Reported
All Persons 14 years and older...	4,043	2,430	254	328	642	389
Pair 1:						
Enumerator A....	480	312	28	24	96	20
Enumerator B....	463	290	15	67	22	69
Pair 2:						
Enumerator A....	841	530	86	50	91	84
Enumerator B....	790	395	55	103	158	79
Pair 3:						
Enumerator A....	740	500	28	37	169	6
Enumerator B....	729	403	42	47	106	131

response bias, B. When the reinterview can be considered as providing generally more accurate responses, the estimator of B is the net difference rate (*NDR*). However, *NDR* is often computed for items for which the reinterview is considered as an identical replication of the census interview. In these cases, *NDR* is not an estimator of B but is used as an indicator of how well the reinterview replicated the census procedure. Thus, *NDR* is usually reported regardless of the objectives of the reinterview survey. However, one must be cautious in its interpretation.

Figure 3-13 also describes the computations for the index of inconsistency, I, and an aggregate measure of inconsistency, I_{AG}. The latter measure may be regarded as a weighted average of indexes of inconsistency across all categories for the item. The interpretation of I_{AG} for the item is the same as the interpretation of I for each category.

Computational formulas for the standard errors for the estimator are provided in figure 3-14.

3.2 *Reporting the results*

Statistical offices produce a vast range of statistics from a census, which are employed by users for a variety of purposes. All the purposes for which the census data or the census evaluation study results will be used are not always known in advance. Therefore, it is important for a statistical office to provide its users with information on the quality of the data to assist the users in interpreting the results and in deciding whether to use them for a given purpose.

The following two types of reports on the results of census evaluation studies may be distinguished: (a) reports with a more technical orientation that are primarily prepared for professional statisticians and other individuals of a similar background, and (b) reports with a less technical orientation which are prepared for users of the census data. The guidelines set out below refer in some degree to both types of audiences but primarily to the latter group.

Figure 3-11. GENERAL NOTATION FOR COMPUTING RESPONSE ERROR MEASURES IN REINTERVIEW STUDIES

Reinterview Classification (i = 1, 2, ...c)	Total reporting	Classification Reported in Census (j = 1, 2, ... c) Category 1	Category 2	Category j	Category c
Total reporting[1].....	n	$Y_{\cdot 1}$	$Y_{\cdot 2}$	$Y_{\cdot j}$	$Y_{\cdot c}$
Category 1............	$Y_{1\cdot}$	Y_{11}	Y_{12}	Y_{1j}	Y_{1c}
Category 2............	$Y_{2\cdot}$	Y_{21}	Y_{22}	Y_{2j}	Y_{2c}
(etc.)					
Category i............	$Y_{i\cdot}$	Y_{i1}	Y_{i2}	Y_{ij}	Y_{ic}
(etc.)					
Category c............	$Y_{c\cdot}$	Y_{c1}	Y_{c2}	Y_{cj}	Y_{cc}

[1]*This table excludes all cases for which there was no report in either the census, the reinterview, or both.*

Figure 3-12. EXAMPLE OF PROCEDURE FOR COMPUTING RESPONSE ERROR MEASURES

(Table contains artificial data for year the structure was built. Figures underlined denote exact match.)

Reinterview Classification (i = 1, 2, ... 6)	Total Reporting	Classification Reported in Census (j = 1, 2, ... c) 1979 to 1980	1975 to 1978	1970 to 1974	1960 to 1969	1950 to 1959	1949 or earlier	Unreported
Total reporting..	1,218	43	120	159	313	172	411	56
1979 to 1980........	38	<u>27</u>	6	1	3	0	1	5
1975 to 1978........	122	11	<u>85</u>	14	8	0	4	2
1970 to 1974........	160	2	20	<u>111</u>	18	0	4	13
1960 to 1969........	294	0	4	27	<u>237</u>	12	14	12
1950 to 1959........	147	0	1	0	27	<u>95</u>	24	4
1949 or earlier.....	457	3	4	6	20	60	<u>364</u>	20
Unreported.......	43	3	2	10	8	3	17	<u>8</u>

In order to meet the users' basic needs for indications of the scope and applicability of the evaluation study, and to help the user to apply the evaluation study results to determine the quality of the census data, the evaluation report generally should incorporate the structure and provide the information described below.

3.21 <u>Introduction</u>.--The objectives of the evaluation study and the general approach used to accomplish the objectives (that is, content reinterview, interpenetration study, record check study, etc.) should be described in non-technical and easily understood language. The contents and structure of the report also should be described.

3.22 <u>Data source and definitions</u>.--Basic information should be provided on the mode of interviewing (that is, telephone, mail, personal interviewing) for both the census and the evaluation study; and all important concepts and technical terms used in the

Figure 3-13. COMPUTING NET DIFFERENCE RATE AND INDEX OF INCONSISTENCY

Net difference rate for category i (an estimator of B (response bias) only when the reinterview response is considered to be the "truth"):

$$NDR = \frac{(Y_{.i} - Y_{i.})}{n} \times (100), \quad (i=1, \ldots, C)$$

Net difference rate, year-built interval "1950 to 1959":

$$NDR = \frac{172 - 147}{1{,}218} \times (100) = \frac{25}{1{,}218} \times (100) = 2.05\%$$

Index of inconsistency for category i (appropriate only when the reinterview response is considered to be in replication of the census response):

$$\hat{I} = \frac{(Y_{.i} + Y_{i.} - 2Y_{ii})}{\frac{1}{n}\left[Y_{.i}(n-Y_{i.}) + Y_{i.}(n-Y_{.i})\right]} \times (100), \quad (i=1, \ldots, C)$$

NOTE: Y_{ii} is the i^{th} diagonal term.

Index of inconsistency for year-built interval "1950-1959":

$$\hat{I} = \frac{[172 + 147 - 2(95)]}{\frac{1}{1{,}218}\left[172(1{,}218 - 147) + 147(1{,}218 - 172)\right]} \times (100)$$

$$= \frac{319 - 190}{\frac{1}{1{,}218}\left[172(1{,}071 + 147(1{,}046)\right]} \times (100) = \frac{129}{277.48} \times (100)$$

$$= 46.49$$

Aggregate index of inconsistency (appropriate only when the reinterview response is considered to be a replication of the census response):

$$\hat{I}_{AG} = \frac{\left(n - \sum_{i}^{C} Y_{ii}\right)}{\left(n - \frac{1}{n}\sum_{i}^{C} Y_{.i}Y_{i.}\right)} \times (100)$$

Aggregate index of inconsistency, year built:

$$I_{AG} = \frac{1{,}218 - 919}{1{,}218 - \frac{1}{1{,}218}\left[(43)(38)+(120)(122)+(159)(160)+(313)(294)+(172)(147)+(411)(457)\right]} \times (100)$$

$$= \frac{1{,}218 - 919}{1{,}218 - \left(\frac{1}{1{,}218}\right)(346{,}847)} \times (100) = 32\%$$

Figure 3-14. COMPUTING STANDARD ERRORS AND NINETY-FIVE PERCENT CONFIDENCE INTERVALS

Ninety-five percent confidence interval of net difference rate for category i:

$$(i=1, \ldots, C)$$

Ninety-five percent confidence limits are:

$$\frac{(Y_{.i} - Y_{i.}) \pm 2\sqrt{Y_{.i} + Y_{i.} - 2Y_{ii} + 1}}{n} \times (100)$$

Exceptions:

(1) If $(Y_{i.} - Y_{ii}) = 0$, then widen the high ninety-five percent confidence limit by adding:

$$\left[\frac{2}{n} \times (100)\right]$$

(2) If $(Y_{.i} - Y_{ii}) = 0$, then widen the low ninety-five percent confidence limit by subtracting:

$$\left[\frac{2}{n} \times (100)\right]$$

(3) If both (1) and (2) above, the ninety-five percent confidence limits are estimated as:

$$\left[\frac{-4}{n} \times (100)\right] \text{ to } \left[\frac{+4}{n} \times (100)\right]$$

Ninety-five percent confidence interval of net difference rate for year-built interval "1950-1959":

(1) Low ninety-five percent confidence limit is:

$$\frac{(172 - 147) - 2\sqrt{172 + 147 - 2(95) + 1}}{1{,}218} \times (100)$$

$$\frac{25 - 2(11.40)}{1{,}218} \times (100) = 0.18$$

(2) High ninety-five percent confidence limit is:

$$\frac{(172 - 147) + 2\sqrt{172 + 147 - 2(95) + 1}}{1{,}218} \times (100)$$

$$= \frac{25 + 2(11.40)}{1{,}218} \times (100) = 3.92$$

Figure 3-14. COMPUTING STANDARD ERRORS AND NINETY-FIVE PERCENT CONFIDENCE INTERVALS--(continued)

Ninety-five percent confidence interval of index of inconsistency for category i:

$$(i=1, \ldots, C)$$

(1) If $\left(\frac{Y_{.i} + Y_{i.} - 2Y_{ii}}{n}\right) \leq .10$, ninety-five percent confidence limits are:

$$\frac{(Y_{.i} + Y_{i.} - 2Y_{ii} + 2) \pm 2\sqrt{Y_{.i} + Y_{i.} - 2Y_{ii} + 1}}{Y_{.i}\left(1 - \frac{Y_{i.}}{n}\right) + Y_{i.}\left(1 - \frac{Y_{.i}}{n}\right)} \times (100)$$

(2) If $\left(\frac{Y_{.i} + Y_{i.} - 2Y_{ii}}{n}\right) > .10$, ninety-five percent confidence limits are:

$$\frac{(Y_{.i} + Y_{i.} - 2Y_{ii} + 2) \pm 2\sqrt{\frac{1}{n}(Y_{.i} + Y_{i.} - 2Y_{ii})(n_{..} - Y_{.i} - Y_{i.} + 2Y_{ii})}}{Y_{.i}\left(1 - \frac{Y_{i.}}{n}\right) + Y_{i.}\left(1 - \frac{Y_{.i}}{n}\right)} \times (100)$$

Ninety-five percent confidence interval of index of inconsistency for year-built interval "1950-1959":

(1) $\frac{172 + 147 - 2(95)}{1,218} = .106$

(2) Low ninety-five percent confidence limit is:

$$\frac{\{172 + 147 - 2(95) + 2\} - 2\sqrt{\frac{1}{1,218}\{(172+147) - 2(95)\}\{1,218-172-147+2(95)\}}}{172\left(1 - \frac{147}{1,218}\right) + 147\left(1 - \frac{172}{1,218}\right)} \times (100)$$

$$= \frac{131 - 2\sqrt{115.337}}{151.24 + 126.24} \times (100) = \frac{131 - 21.48}{277.48} \times (100) = 39.47$$

(3) High ninety-five percent confidence limit is:

$$\frac{\{172 + 147 - 2(95) + 2\} + 2\sqrt{\frac{1}{1,218}\{(172+147) - 2(95)\}\{1,218-172-147+2(95)\}}}{172\left(1 - \frac{147}{1,218}\right) + 147\left(1 - \frac{172}{1,218}\right)} \times (100)$$

$$= \frac{131 + 21.48}{277.48} \times (100) = 54.95$$

Figure 3-14. COMPUTING STANDARD ERRORS AND NINETY-FIVE PERCENT CONFIDENCE INTERVALS--(continued)

Ninety-five percent confidence interval for the aggregate index of inconsistency:

(1) If $\left[\frac{n - \sum_{i=1}^{C} Y_{ii}}{n}\right] \leq .10$, ninety-five percent confidence limits are:

$$\frac{\left(n - \sum_{i=1}^{C} Y_{ii} + 2\right) \pm 2\sqrt{n - \sum_{i=1}^{C} Y_{ii} + 1}}{\left(n - \frac{1}{n} \sum_{i=1}^{C} Y_{.i}Y_{i.}\right)} \times (100)$$

(2) If $\left[\frac{n - \sum_{i=1}^{C} Y_{ii}}{n}\right] > .10$, ninety-five percent confidence limits are:

$$\frac{\left(n - \sum_{i=1}^{C} Y_{ii} + 2\right) \pm 2\sqrt{\frac{1}{n}\left(n - \sum_{i=1}^{C} Y_{ii}\right)\left(\sum_{i=1}^{C} Y_{ii}\right)}}{\left(n - \frac{1}{n} \sum_{i=1}^{C} Y_{.i}Y_{i.}\right)} \times (100)$$

Ninety-five percent confidence interval for the aggregate index of inconsistency, year built:

(1) $\frac{1,218 - (27+85+111+237+95+364)}{1,218} = .25$

(2) Low ninety-five percent confidence limit is:

$$\frac{1,218-(27+85+111+237+95+364)+2-2\sqrt{\frac{1}{1,218}\{1,218-(27+85+111+237+95+364\}\{27+85+111+237+95+364\}}}{1,218 - \frac{1}{1,218}\{(43)(38)+(120)(122)+(159)(160)+(313)(294)+(172)(147)+(411)(457)\}} \times (100)$$

$$= \frac{1,220 - 919 - 2\sqrt{225.60}}{1,218 - \frac{346,847}{1,218}} \times (100) = \frac{301 - 2(15.02)}{933.23} \times (100) = 29.03$$

(3) High ninety-five percent confidence limit is:

$$\frac{1,218-(27+85+111+237+95+364)+2+2\sqrt{\frac{1}{1,218}\{1,218-(27+85+111+237+95+364\}\{27+85+111+237+95+364\}}}{1,218 - \frac{1}{1,218}\{(43)(38)+(120)(122)+(159)(160)+(313)(294)+(172)(147)+(411)(457)\}} \times (100)$$

$$= \frac{301 + 2(15.02)}{933.23} \times (100) = 35.47$$

report should be explained in a clear and unambiguous manner. For example, the correlated component of enumerator variance, index of inconsistency, response bias, or any other technical terms used in the report should be defined. There should also be a clear discussion of how the reported measures of response error are interpreted. In addition, the date and time period of the study in relation to the census should be specified.

3.23 Field procedures, study design, and estimation methods.--The essential aspects of the field procedures, study design, and estimation methods should be discussed. Any unforeseen, unusual, or atypical events that occurred which might significantly affect the validity of the evaluation results (violations of interpenetration procedures, misunderstandings on the part of the enumerators or respondents, strong negative reactions of the field staff or the public to the evaluation) should be specifically mentioned. Descriptions of the clerical and computer processing procedures should also be included. For reinterview procedures, describe how the field collection and data processing procedures differed from the census and how this may affect the results.

3.24 Results for housing characteristics.--Tables giving the results of the evaluation study along with discussions of the important findings for each housing characteristic should be presented. The following outline may be used:

(1) Title line giving the name of the characteristic (that is, age, marital status, etc.).

(2) Verbatim statement of the question as it appeared on the census questionnaire and, for reinterview studies, on the reinterview questionnaire.

(3) For reinterview studies only, a detailed table of the results using the format of figure 3-15 (this is essentially the same as figure 3-11) state whether the reinterview objective for the question was to measure simple response variance or response bias. If the objective was response bias, discuss the efforts made to obtain the truthful response and how well these efforts may have succeeded.

(4) For reinterview studies, a summary table of the form in figure 3-16 should be used.

(5) For interpenetration studies of the enumerator variance, provide a summary table of the form in figure 3-17.

(6) Discuss the results, indicating whether the reported measures of response error are particularly large or small (see section 3.3). Other analyses may be provided to investigate the sources of large indices of inconsistencies, response biases, or intra-interviewer correlation coefficients. For example, measures may be reported, using table formats similar to figure 3-16, for age-race-sex subgroups and other cross-classified tables.

3.25 Results for population characteristics.--Tables giving the results of the evaluation study for population characteristics should be presented. The outline and structure for the section is identical to those previously given for housing characteristics.

3.26 Data limitations.--This section should discuss all types and sources of error, problems in the fieldwork or design of the study, or difficulties in the processing of the data that are considered to have an important bearing on the quality, applicability, or limitations of the data. The principal types of limitations to be considered in this regard are, for reinterview studies:

(1) Matching errors and problems.

(2) Reinterview nonresponse rates.

(3) Violations in the assumptions of the replication reinterview or the improved reinterview procedure, whichever was the objective.

(4) The population sampled for the reinterview and its correspondence to the census population.

Figure 3-15. EXAMPLE OF A DETAILED REINTERVIEW TABLE

Name of Characteristic

Reinterview Classification	Total matched housing units	Classification Reported in Census ($j = 1, 2, \ldots c$) Total Reporting	Category 1	Category 2		Category c
Total matched housing units..	n'	n_C	$Y'_{\cdot 1}$	$Y'_{\cdot 2}$		$Y'_{\cdot c}$
Total reporting......	n_R	n	$Y_{\cdot 1}$	$Y_{\cdot 2}$		$Y_{\cdot c}$
Category 1..........	$Y'_{1\cdot}$	$Y_{1\cdot}$	Y_{11}	Y_{12}		Y_{1c}
Category 2..........	$Y'_{2\cdot}$	$Y_{2\cdot}$	Y_{21}	Y_{22}		Y_{2c}
(etc.)						
Category c..........	$Y'_{c\cdot}$	$Y_{c\cdot}$	Y_{c1}	Y_{c2}		Y_{cc}

Note: The notation in this table is identical to that of figure 3-11 with the exception that column 1 and row 1 have been added. For this column and row:

n' *is the total number of sample cases, including nonresponses in either the census or reinterview.*

n_C *is the total number of sample cases reporting in the census. Thus* $Y'_{\cdot i}$ *is the number of these* n_C *cases reporting in category* i *in the census.*

n_R *is the total number of sample cases reporting in the reinterview. Thus* $Y'_{j\cdot}$ *is the number of these* n_R *cases reporting in category* j *in the reinterview.*

Figure 3-16. EXAMPLE OF A SUMMARY TABLE FOR REPORTING MEASURES FROM REINTERVIEW STUDIES

Name of Characteristic

Classification	Total matched housing units	Percent of total	Index of inconsistency I	Lower confidence limit	Upper confidence limit	Net difference rate	Lower confidence limit	Upper confidence limit
Total matched housing units..		100	(x)	(x)	(x)	(x)	(x)	(x)
Category 1......								
Category 2......								
(etc.)								
Category c......								

(x) Not applicable

Figure 3-17. EXAMPLE OF SUMMARY TABLE FOR REPORTING MEASURES OF ENUMERATOR VARIANCE

Name of Characteristic

Characteristic	Total housing units	Percent of total	Correlated component $\hat{CC}$	Lower confidence limit	Upper confidence limit	Intra-interview correlation coefficient ρ_y	Lower confidence limit	Upper confidence limit
Total housing units								
Category 1.....								
Category 2.....								
(etc.)								
Category *c*.....								

3.27 Appendices.--The basic formulas used in the computations of the reported response error measures and their variances should be documented in the appendices to the report.

3.3 *Interpretation of the results*

As a general rule for interpreting the magnitudes of the response error measures, figure 3-18 may be useful for the analysis. It is based on the experiences in the United States with evaluation studies conducted for personal interviewing in censuses and surveys.

4. PROFILES OF INTERNATIONAL CONTENT EVALUATION STUDIES

In this section, some examples of content evaluation studies are described, and the techniques of the previous sections are illustrated.

4.1 *Bangladesh*

As part of the post-enumeration check (PEC) survey of the 1981 census of Bangladesh, a reinterview study with the "response variance objective" was conducted.

The sample was a systematic sample of 250 enumeration areas (EA's) selected from the 211,751 EA's included in the census. Before sampling, the EA's were sorted by urban/rural classification and geographic location. The average size of the EA's was about 72 housing units. The final sample size was 20,962 households or 121,072 persons.

In striving to maintain independence between the PEC and the census, reinterview enumerators were especially recruited for the job independently of the census. The reinterview survey was conducted 2 1/2 weeks after the census.

Following the survey, each EA was independently matched twice to the census with a third party adjudicating discrepancies. This yielded 104,703 matched persons for the content analysis.

Figure 3-19 summarizes the results of the study. As expected, sex was the most reliably reported characteristic (I_{AG}=0), and age was the least reliable with I_{AG} in the moderate range. An interesting finding from the survey was that reliability of reporting for age decreased steadily as age increased.

Note that, from the table, the "rate of agreement" between the census and the reinterview (1 - the gross error rate) is sometimes smaller for some characteristics which also

Figure 3-18. GENERAL GUIDELINES FOR INTERPRETING THE MAGNITUDES OF RESPONSE ERROR MEASURES FOR CENSUSES

Response Error Measure	Low Range	Moderate Range	High Range
Index of inconsistency, I.....	< 20	20 - 50	> 50
Aggregate index of inconsistency, I_{AG}...	< 20	20 - 50	> 50
Absolute value of the the ratio of the net difference rate to the population mean of proportion, $\|NDR/P\|$.............	<.01	.01-.05	>.05
Intra-interviewer correlation coefficient, ρ_y:			
Simple[1] items......	<.001	.001-.005	>.005
Difficult[2] items...	<.005	.005-.02	>.02

[1] *"Simple items" refer to age, sex, marital status, and other items that are neither ambiguous, emotionally loaded, or subject to memory loss.*

[2] *"Difficult items" refer to education, income, migration, some occupation entries, and other items requiring greater recall and/or effort on the part of the respondent or enumerator.*

have smaller aggregate indexes of inconsistency, such as literacy and tenure. Thus, "rate of agreement" is not a good measure of the relative impact of simple response variance on the total variance for an item. It is, however, a good measure of the gross errors for an item. Coefficients of variation for the estimates were not provided in the report, but are probably in the order of about 1 percent.

4.2 India

As part of the PEC for the 1981 census of India, a reinterview survey was conducted. The objective of the survey was to determine the reliability of reporting age, literacy, and other variables related to labor force status. Thus, the reinterview survey had the response variance objective.

Figure 3-19. AGGREGATE INDEX OF INCONSISTENCY FOR SELECTED CHARACTERISTICS: BANGLADESH

Characteristics	I_{AG}	Rate of Agreement (percent)
Sex................	0.0	99.9
Marital status.....	5.7	97.0
Literacy...........	18.1	94.0
Type of roof material.........	21.1	89.0
Type of wall material.........	27.7	85.0
Educational attainment.......	28.6	88.0
Tenure.............	29.8	95.0
Age................	41.2	63.0

The reinterview survey was conducted 2 weeks after the census in a subsample of 750 of the 4,000 PEC blocks. In rural areas, 546 blocks were selected and, in urban areas, 204 were selected. Each block averaged about 122 housing units. Thus, the total sample size for the survey was about 91,500 households.

Figure 3-20 is indicative of the type of analyses that were conducted. This table gives the index of inconsistency and the net difference rate for literacy and age by sex and rural/urban area. Similar analyses were performed for the labor force variables. The net difference rate analyses may provide an approximation to the level of under- or over-reporting, to the extent that the reinterview responses for age were closer to the "truth" than the census responses (i.e., reinterview misclassification error rates were small). Since the reinterview enumerators were more carefully selected and better trained and supervised, this premise seems plausible. However, no attempt was made to improve the questioning of age, nor was there a reconciliation of discrepancies between the census and the reinterview.

To provide additional information on the reliability of the reporting of age, the reinterview enumerator was asked to record his/her

Figure 3-20. RESPONSE VARIANCE AND RESPONSE BIAS BY SEX, LITERACY, AND AGE

Characteristic	Index of inconsistency		Net difference rate		Enumerator assessment of reliability of age (percent)	
	Male	Female	Male	Female	Male	Female
LITERACY						
Urban	7.71	7.71	.11	-.57	(x)	(x)
Rural	9.19	9.39	.47	-.26	(x)	(x)
AGE						
0 to 4 years:						
Urban	13.75	17.35	.52	-.01	72.61	70.30
Rural	7.97	5.24	.20	.17	68.80	72.37
5 to 9 years:						
Urban	14.73	16.04	.21	-.22	68.31	71.33
Rural	9.44	8.65	.14	-.95	66.15	66.90
20 to 24 years:						
Urban	15.99	18.27	-.34	-.79	77.66	66.52
Rural	21.95	28.40	-.41	-.25	65.00	58.31
25 to 29 years:						
Urban	19.94	22.61	.19	-.17	67.84	61.59
Rural	23.80	26.11	.14	.08	61.74	57.75
45 to 49 years:						
Urban	24.92	29.76	.66	.23	71.43	61.90
Rural	31.06	34.06	-.31	-.39	59.03	52.97
50 to 54 years:						
Urban	31.51	23.00	-.15	.08	50.55	63.57
Rural	31.80	28.62	-.13	.38	60.30	69.96
65 to 69 years:						
Urban	38.80	34.77	-.37	.33	51.75	57.25
Rural	38.80	41.71	-.13	-.17	46.74	47.70
70 years or older:						
Urban	16.81	19.66	.00	.03	55.27	50.46
Rural	14.50	16.72	.21	.00	46.59	49.19

(x) Not applicable

opinion as to whether the recorded age was reliable. The percent of responses judged as reliable also appears in figure 3-20. There is evidence of a strong correlation between the enumerator assessment and the index of inconsistency. Standard errors for the estimates are not available.

4.3 *United States*

Content evaluations have been a part of the U.S. census since 1950. Always including reinterview surveys, the censuses have also included interpenetration studies for the study of enumerator variance (Hanson and Marks 1958; U.S. Bureau of the Census, Evaluation: Effects of Interviewers and Crew Leaders, 1960; and U.S Bureau of the Census 1979), coder variance (in 1960), supervisor variance (U.S. Bureau of the Census, Evaluation: Effects of Interviewers and Crew Leaders, 1960), and editing and telephone followup enumerator variance (Biemer and Katzoff 1980).

The last census to be conducted completely by face-to-face interviewing was the 1950 census of population. It was as a result of the enumerator variance study conducted for that census that the U.S. Census Bureau abandoned personal interviewing for censuses in favor of self-enumeration using a mail-out/

mail-back approach. In this section, the result of that study will be briefly described.

The 1950 Enumerator Variance Study (EVS) was conducted in 21 purposively selected counties in Ohio and Michigan. A total of 200 strata with an average population of about 6,500 persons each, were formed. Within each stratum, enumeration districts were paired at random, and each pair was assigned to the interviewers by a random method. See (Hanson and Marks 1958) for a complete description of the 1950 EVS.

Figure 3-21 shows the estimates of ρ_y for selected 1950 population characteristics. The variances of the estimates were not computed. Furthermore, the estimates may not be representative of the population of 1950 census enumerators, since the study was concentrated in 21 counties which were very similar to each other rather than spread over the entire United States.

As is obvious from the figure, most of the items exhibit enumerator effects which are in the moderate-to-large range. The non-response categories consistently showed large ρ_y's indicating that enumerators have a considerable effect on item nonresponse rates.

For the 1960 census, another enumerator interpenetration study was conducted, referred to as the Response Variance Study (RVS). A major objective of the RVS was to evaluate the reduction in census variance as a result of the change to self-enumeration. About 45 percent of all census questionnaires were completed solely by respondents in 1960.

A number of design differences between the 1950 EVS and the 1960 RVS limit the comparisons. However, the RVS concluded that enumerator effects were dramatically reduced largely as a result of self-enumeration. For example, for educational attainment the reduction was about 80 percent and for income about 50 percent. Overall, enumerator effects were reduced by about 75 percent. The items showing the least reduction were the nonresponse items which, as in 1950, had the highest enumerator effects. The RVS analyses also revealed that enumerator effects on item nonresponse vary considerably by geographic area.

The 1950 and 1960 enumerator variance studies are illustrations of the usefulness of response variance analysis for evaluating alternative census design choices.

Figure 3-21. VALUES OF ρ_y FOR SELECTED CHARACTERISTICS, U.S. POPULATION: 1950

Characteristic	ρ_y
Race:	
Negro	.0165
Other	.0043
Age:	
Under 1 year	.0002
1 year or older	.0009
Under 14 years	.0030
25 years or older	.0040
55 years or older	.0026
Educational attainment:	
Grade 12 or over	.0125
Grade 9 or over	.0064
Grade 8 or over	.0064
Grade 5 or over	.0027
Not reported	.0543
Residence:	
Farm	.0609
Nonfarm:	
Male	.0330
Female	.0278
Income--earned:	
Less than 2000	.0059
2000 to 4999	.0060
5000 or more	.0087
Not reported	.0160
Income--unearned:	
None	.0313
Less than 2500	.0246
2500 or more	.0009
Not reported	.0599

Source: 1950 U.S. Population Census

5. *REFERENCES*

Bailar, B.A. 1968. "Recent Research in Reinterview Procedures." *Journal of the American Statistical Association* 63:41-63.

-----. 1976. "Some Sources of Error and Their Effects on Census Statistics." *Demography* 13:273-296.

Bailar, B.A., and Biemer, P.P. 1984. "Some Methods for Evaluating Nonsampling Error in Household Censuses and Surveys." In *W.G. Cochran's Impact on Statistics*, edited by P.S.R.S. Roo, 253-274. New York: John Wiley and Sons, Inc.

Bangladesh Bureau of Statistics. 1983. *1981 Bangladesh Population Census: Post-Enumeration Check*. Dakha: Bangladesh Bureau of Statistics.

Biemer, P.P., and Katzoff, E.B. 1980. "Estimation of Nonsampling Error Due to Selected Office Operations of the 1980 Census." In *Proceedings of the American Statistical Association, Section of Survey Research Methods*. Houston: American Statistical Association.

Biemer, P.P., and Stokes, S.L. 1985. "Optimal Design of Interviewer Variance Experiments in Complex Surveys." *Journal of the American Statistical Association* 80:158-166.

Hansen, M.H.; Hurwitz, W.N.; and Berabod, M.A. 1959. "Measurement Errors in Censuses and Surveys." *Proceedings of the American Statistical Association, Section of Social Statistics*. Washington, D.C.: American Statistical Association.

Hansen, M.H.; Hurwitz, W.N.; and Pritzker, L. 1964. "The Estimation and Interpretation of Gross Differences and the Simple Response Variance." In *Contributions to Statistics*, edited by C.R. Rao, 111-136. Oxford, Pergamon Press.

Hanson, R.H., and Marks, E.S. 1958. "Influence of the Interviewer on the Accuracy of Survey Results." *Journal of the American Statistical Association* 53:635-655.

Indian Administrative Service. 1982. *Census of India 1981: Report on the Post-Enumeration Check*. Series 1, No. 4, New Delhi: Indian Administrative Service.

Kish, R. 1962. "Studies of Interviewer Variance for Attitudinal Variables." *Journal of the American Statistical Association* 57:92-115.

Marks, E.S.; Mauldin, W.P.; and Nisselson, H. 1953. "The Post-Enumeration Survey of the 1950 Census: A Case History in Survey Design." *Journal of the American Statistical Association* 48:220-243.

U.S. Bureau of the Census. 1960. *The Post-Enumeration Survey: 1950*. Technical Paper No. 4. Washington, D. C.: U.S. Bureau of the Census.

-----. 1963. *Evaluation and Research Program of the U.S. Censuses of Population and Housing, 1960: Background, Procedures, and Forms*. Series ER 60, No. 1. Washington, D.C.: U.S. Bureau of the Census.

-----. 1968. *Evaluation and Research Program of the U.S. Censuses of Population and Housing, 1960: Effects of Interviewers and Crew Leaders*. Series ER 60, No. 7. Washington, D.C.: U.S. Bureau of the Census.

-----. 1975. *Accuracy of Data for Selected Housing Characteristics as Measured by Reinterviews*. Series PHC(E)-10. Washington, D.C.: U.S. Bureau of the Census.

-----. 1975. *Accuracy of Data for Selected Population Characteristics as Measured by Reinterviews*. Series PCH(E)-9. Washington, D.C.: U.S. Bureau of the Census.

-----. 1979. *Enumerator Variance in the 1970 Census*. Series PHC(E)-13. Washington, D.C.: U.S. Bureau of the Census.

Chapter 4. OVERVIEW TO CENSUS EVALUATION THROUGH DEMOGRAPHIC ANALYSIS

1. INTRODUCTION

Demographic methods play a very important role in the census evaluation programs in most, if not all, countries. In countries where matching studies have been conducted as part of the evaluation program, demographic methods may be used to supplement the results of these studies in order to provide further insight into the nature and magnitude of various types of errors in the data. In cases where the implementation of a PES or one of the other types of matching studies described in chapters 2 and 3 are not feasible, demographic and/or related non-matching techniques provide the only basis for assessing the quality of a census enumeration.

Depending upon the amount of demographic information available about a country, demographic analyses can provide considerable insight into the magnitude and nature of errors in census data. In the most limited case where only the results of a single census are available, it is usually possible to recognize (at minimum) whether some combination of coverage and content error has resulted in an implausible enumeration of one or more segments of the population. In countries in which two or more censuses have been conducted, a somewhat wider range of methods is applicable and defensible estimates of the magnitude of census coverage and content error can often be derived without the benefit of a PES. The availability of results from one or more demographic surveys conducted during the intercensal period further increases the usefulness of demographic analyses. The substitutability of demographic analyses for a PES or other matching study, particularly with regard to the question of the completeness of census coverage, will in general be determined by the amount of demographic information available about the country and the extent to which the composition of the country's population has been affected by extraordinary events (wars, famines, major epidemics, etc.) and other distorting factors (abrupt changes in fertility or mortality levels or significant levels of net international migration, for example).

This chapter is the first of two chapters describing the use of demographic and related (non-matching) methods for the evaluation of population censuses. The present chapter provides an overview to and background for the detailed description and illustrative applications of selected demographic methods presented in chapter 5. Specifically, the purposes of this chapter are: (a) to outline the underlying rationale of the demographic approach to census evaluation, (b) to briefly describe methods having practical applications for census evaluation purposes, and (c) to summarize techniques available for obtaining indirect estimates of demographic parameters required for the application of some of the more useful demographic methods of census evaluation.

The rationale for including indirect estimation techniques in this manual is that since estimates of selected parameters are required in the application of a number of the more powerful demographic techniques for census evaluation and many countries lack registration data of sufficient quality to produce accurate direct estimates, it will often be the case that the indirect estimation of one or more basic demographic parameters will have to be undertaken as part of the census evaluation process. Accordingly, a summary of the methods available for estimating demographic parameters in the absence of directly usable registration data is included in the manual to assist in making optimal use of the demographic information available from surveys or previous censuses in developing parameter estimates. The reader is referred to United Nations (1967 and 1983) and U.S. Bureau of the Census (1975) for a more comprehensive treatment of indirect estimation techniques. The uses of these estimates for census evaluation purposes are described and illustrated in chapter 5.

2. *DEMOGRAPHIC METHODS FOR CENSUS EVALUATION*

The fundamental problem to be addressed in evaluating a census in the absence of data suitable for a matching study is one of deriving an "expected" population against which to compare the results of the census under evaluation. Under the demographic approach to census evaluation, knowledge of regularities in the structure of human populations is combined with data on and/or assumptions about demographic parameters in a particular population to derive an estimate as to what the size and composition of the population was likely to have been at the reference date of the census to be evaluated. This "expected" population is then used as a standard by which to assess the quality of the census enumeration.

The methods described in this chapter represent alternative methods and procedures for deriving a "standard" population for evaluation purposes, contingent upon the types of demographic information available about a particular population, the quality (accuracy) of that information, and the assumptions about levels and trends in the components of population growth which can be justified in the particular case. These criteria are relevant both for the choice of evaluation methodology and for decisions as to which estimate of census error to accept as the most plausible in instances where estimates can be derived on the basis of more than one method.

In the presentation below, the various methods considered are divided into three categories on the basis of their data requirements and underlying approach as follows: (a) methods based upon direct comparisons with data from external sources; (b) methods based upon comparison with theoretical distributions; and (c) methods based upon comparisons with a previous census or censuses.

2.1 *Direct comparisons with data from external sources*

One of the more conceptually straightforward approaches for assessing the plausibility of a census count involves the comparison of the census count with a corresponding count from an independent data source. Among the possible sources of data for such comparisons are population registers, vital registration systems, baptismal records, school enrollment data, citizen identification systems, social security/health systems, and existing household surveys.

The basic assumption underlying the use of such data for census evaluation purposes is that the count of persons derived from the external source is at least as accurate as the corresponding count or distribution obtained from the census. To the extent that this assumption is justified, the direct comparison of census counts with these "expected" counts will provide valuable information on the accuracy of the census enumeration. Incomplete and/or defective data limit the usefulness of the comparisons, but in some instances such data may nevertheless provide useful information for census evaluation purposes. For example, if a justifiable estimate of the level of undercoverage in the particular statistical program being compared with the census is available or can be derived, the "corrected" total can be used to provide a rough estimate of the completeness of census coverage. Even when such adjustments are not possible, data from external data sources may be used to estimate the "lower bound" of the completeness of coverage in a census. In most practical census evaluation applications, it will be necessary to make some form of adjustment to the data from the external data source to compensate for such shortcomings as incomplete population coverage, different target populations in comparison with the census, differences in reference periods, etc.

The most comprehensive source of information for evaluating a census on a direct-comparison basis is a population register, in which all changes in population resulting from births, deaths, and migration (internal and international) are recorded. In countries in which population registers are maintained, registration is usually undertaken for small geographic areas, making it possible to evaluate a census enumeration with a high degree of geographical detail. In addition, the detailed information collected in the register provides a basis for a detailed evaluation of both the coverage of the census and the accuracy of the recorded characteristics of the population. Unfortunately, only a few countries maintain complete population registers.

A more commonly available source of information for census evaluation purposes is a vital registration system. A considerable number of census evaluations have been undertaken in which the registered number of births in the months/years immediately preceding the census (adjusted for infant and child mortality) have been used as a basis for assessing census coverage of the population at the youngest ages (which are frequently under-enumerated in both developed and developing countries). Birth and death registration data may also be used in conjunction with data from two successive censuses. In this application, the registered numbers of births and deaths during the intercensal period (adjusted for undercoverage as necessary and feasible) are added to or subtracted from the population enumerated in the first census to derive an expected population at the reference date of the census being evaluated (see section 6.3 in chapter 5 for further elaboration of this procedure).

Data from baptism records have been used on occasion as a proxy for birth registration in a number of historical demographic studies. In cases where a large proportion of births are covered in such record systems, these may be used as a substitute for birth registration data in one or more of the procedures described above. In general, however, there appear to be relatively few instances of usable record systems of this type in contemporary developing countries.

In countries in which the proportion of children and adolescents who are enrolled in school is high, school enrollment data can be used to evaluate census coverage of the school age population. A number of studies have been conducted in developing countries using such data for census evaluation purposes. In a recent study of the Soviet Union, for example, Anderson and Silver (1985) used primary school enrollment data to evaluate the 1979 Soviet census. They concluded that these data provided a more appropriate standard for assessing the completeness of coverage of Soviet censuses than birth registration data due to their higher degree of completeness.

In addition to the problem of less than complete school records of children of school age, the use of school enrollment data for evaluating censuses is limited in many countries by factors such as variation in the quantity and quality of information collected by type of school (public versus private, for example) and by geographical area, as well as by variations in the ages at which children enter and leave school. In some instances, however, it may be possible to overcome these problems for at least portions of a country (in urban areas, for example) and derive a useful standard for census evaluation purposes.

In countries which have social security or similar old-age security systems in which a significant proportion of the population above a particular age (age 65, for example) is registered, these data can be used to evaluate the coverage of this segment of the population. The general evaluation approach is similar to that described above in connection with other registration systems, as are the major limitations. In actual practice, few registration systems are fully complete and thus some basis will usually be needed by which to adjust the registration data for under-registration of the target population and other incompatibilities with the census.

The external sources of data cited above are illustrative of the types of registration data which may be utilized in evaluating the plausibility of census counts for the total or for selected subsets of the population on a direct comparison basis. Similar types of data not mentioned above which are available in particular cases may be used in the manner described above provided that three essential conditions are met: (a) the registration/statistical system is independent of the census; (b) the target population for the system is substantially the same as the population covered by the census, or an important analytic subset of them; and (c) the system is either substantially complete or an approximate estimate of the degree of under- or over-coverage is available. Where these conditions cannot be met, census evaluations must rely upon one or more methods described later in this chapter.

Finally, in countries which have conducted one or more household sample surveys within a few years of the census, it may be possible to use the results of the survey(s) to assess, on a limited basis, the plausibility of the recorded distribution of the census population by selected characteristics. As noted in chapter 1, the rationale for making such comparisons is that because of the greater degree of operational control that can normally be imposed over a survey activity in comparison with a census, survey data are often less seriously affected by nonsampling error. Balanced against this, however, is the fact that survey data are subject to sampling error, which in many cases limits the strength of the conclusions that can be derived from comparisons of survey and census data. Nevertheless, serious discrepancies between census and survey distributions for selected characteristics are indicative of error in one or both sets of data whose cause(s) should be investigated.

Both the utility of and some of the problems inherent in such comparisons for census evaluation pruposes is illustrated in figure 4-1, which shows the distribution of females aged 15 to 49 by current marital status from the 1971 Census of India and a national fertility survey conducted in 1972 (adapted from Bhat et al., 1984). The two distributions are clearly inconsistent, particularly for the youngest and oldest age groups. In this particular case, further investigation suggested that the survey data were flawed by errors made by enumerators in selecting respondents for the survey (Bhat et al., 1984). In other cases, it may be that the survey data will be found to be plausible, from which it would follow that the census data are defective.

Figure 4-1. FEMALE MARITAL DISTRIBUTIONS FROM THE 1971 CENSUS AND 1972 FERTILITY SURVEY: ALL-INDIA

Age group	Percent currently married	
	1971 census	1972 fertility survey
15 to 19 years..	55.4	41.1
20 to 24 years..	88.4	83.7
25 to 29 years..	95.0	92.3
30 to 34 years..	94.1	90.3
35 to 39 years..	91.4	88.6
40 to 44 years..	84.2	86.7
45 to 49 years..	78.2	97.9
10 years or older.........	64.4	60.6

Source: India (1977), reported in Bhat, et. al. (1984).

This general approach can be applied to any census characteristic for which comparable survey estimates are available. An extension of the basic approach would entail the comparison of estimates of demographic parameters (birth or death rates, for example) from census and survey sources. While such comparisons will not provide a basis for a comprehensive evaluation of a census, they often prove useful for identifying census characteristics in which errors are likely and, in some instances, these comparisons provide clues as to the likely causes of the errors.

2.2 *Methods based upon comparison with theoretical distributions*

Methods under this heading take advantage of the fact that, due to biological and cultural factors, the distribution of national populations with respect to several important characteristics (age and sex, for example) follow well-known and fairly predictable patterns for deriving a standard against which the population enumerated in a census can be compared. The age distribution of a population whose fertility has not changed abruptly and has been exposed to at most modest levels of international migration, for example, will normally follow a pattern of diminishing numbers of persons at each successively older age. The distribution of such a population by sex will (in most cases) follow a pattern of an excess of males over females at younger ages, with gradually declining proportions of males at successively older ages. Accordingly, departures from these "expected" distributions which cannot be explained on the basis of distorting factors signify the existence of errors in the census enumeration.

One of the more basic tools for assessing the "reasonableness" of a census age-sex distribution entails the visual inspection of the census counts of population by age and sex using graphical techniques. In the absence of extraordinary events or other distorting factors such as wars, famines, etc., significant swings in fertility or mortality, and substantial levels of net international migration, the age-sex distribution of the census population should resemble the well-known population pyramid (see figure 5-1 for an example of a population pyramid). The exact shape of the pyramid for any particular country is determined by the levels of fertility, mortality, and migration to which it has been exposed, and as such will vary somewhat from country to country. However, the regularities in the age and sex distributions of national populations noted above should nevertheless be apparent from the visual examination of the census age-sex distribution. Significant fluctuations in the size of the population enumerated at different ages and/or for the two sexes at any particular age which cannot be accounted for by one or more of the distorting factors noted above are suggestive of errors in the census enumeration.

The utility of graphic analyses is enhanced somewhat when age-sex distributions from two or more censuses are available. The availability of such data permit the comparison of the enumerated sizes of actual birth cohorts in successive censuses. If affected by mortality alone, the size of birth cohorts should decline systematically in each successive census. Irregularities which cannot be explained on the basis of distorting factors are indicative of error in one or both censuses. A graphical cohort analysis is illustrated in figure 5-11.

The "reasonableness" of a census age-sex distribution can be assessed in a more quantitative fashion through the use of various ratios and indices. These provide numerical measures of the magnitude of departure of a recorded census age-sex distribution from the distribution which would be expected in the absence of census errors and other distorting factors. The ratios and indices convey essentially the same information as may be derived through visual inspection of the data, but may be more useful for analytic purposes due to their quantitative nature.

The more widely used among these ratios are age- and sex-ratios. Age ratios measure the "smoothness" of an age distribution over a restricted age range by means of comparison of the number of persons enumerated at successive ages. If determined by mortality alone, the number of persons at any given age should be approximately equal to the average of the number of persons at the two adjacent ages (the immediately higher and lower ages). To assess the quality of a census enumeration, age ratios as defined in equation (4.1) for 5-year age groups would be calculated for each age group and compared to the expected ratio value of 100. Significant departures from the expected value are indicative either of census error or other distorting factors (or both).

$$(4.1)\qquad {}_5AR_x = \frac{{}_5P_x}{\frac{1}{3}\left({}_5P_{x-5} + {}_5P_x + {}_5P_{x+5}\right)} \times 100$$

Where:

${}_5AR_x$ = The age ratio for the age group x to $x+4$ years

${}_5P_x$ = The population enumerated in the age group x to $x+4$ years

${}_5P_{x-5}$ = The population enumerated in the adjacent younger age group ($x-5$ to $x-1$ years)

${}_5P_{x+5}$ = The population enumerated in the adjacent older age group ($x+5$ to $x+9$ years)

Sex ratios measure the composition of the population enumerated at a particular age with respect to sex. The sex ratio is defined as the number of males per 100 females at any given age x, equation (4.2). In populations which are relatively unaffected by international migration, the expected pattern is for sex ratios to be in the 102 to 107 range at early ages due to a higher proportion of male births. The ratios gradually decline with increasing age of the population due to higher male than female mortality in most (but not all) populations. Significant departures from this expected pattern which cannot be explained by distorting factors (sex-selective migration is a distorting factor which is frequently encountered) are suggestive either of coverage error for population aged x, age-misreporting (content error), or both.

(4.2)

$$_5SR_x = \frac{\text{The enumerated population of males at age } x}{\text{The enumerated population of females at age } x} \times 100$$

The use of age- and sex-ratios is illustrated in section 5.1.

In addition to age- and sex-ratios, a number of summary indices have been developed which measure the magnitude of departures of an observed age-sex distribution from an expected distribution in the form of a single index value. Among the more widely-used of these are the United Nations Age-Sex Accuracy Index, Whipple's Index, and Myers' Blended Index. The United Nations Age-Sex Accuracy Index summarizes the results of age- and sex-ratio analyses into a single score which measures the overall level of distortion in an age-sex distribution (that is, across all ages). The Whipple's and Myers' indices, which require data tabulated by single years of age, measure the magnitude of a particular pattern of error in age distribution - the tendency for ages ending in particular digits (0 and 5, for example) to be reported disproportionately and other ages to be avoided. This pattern of error, known as digit preference or age heaping, is found to some degree in all censuses (see section 5.1 for further discussion of the uses of these indices). Other indices are described in U.S. Bureau of the Census (1975).

An alternative approach for deriving an "expected" age-sex distribution for census evaluation purposes is based upon the application of stable population theory. A stable population is a population which has experienced constant levels of fertility and mortality and no migration over a fairly long period of time (that is, two generations or more). Given a sufficient length of time, the age distribution of such a population will become constant (that is, the proportion of the population at each age x will be unchanging) and will be independent of the initial age distribution.

The proportion of the population at each age in a stable population may be obtained as:

(4.3) $$c(x) = b\,\ell(x)\,\exp\,(-rx)$$

Where:

$c(x)$ = the infinitesimal proportion of the population at exact age x

b = the constant birth rate

r = the constant rate of natural increase

$\ell(x)$ = the probability of survival from birth to age x

The proportion of a stable population under exact age y, $c(y)$, may be expressed as:

(4.4) $$c(y) = \int_0^y b\,\ell(x)\,\exp\,(-rx)dx$$

if w denotes the highest age attainable in the population, it follows that:

(4.5) $$\int_0^w b\,\ell(x)\,\exp\,(-rx)dx = 1.0$$

The practical utility of these relationships is that if two of the parameters, b, r or $\ell(x)$, are known or can be estimated for an actual population which has experienced relatively constant fertility, negligible levels of migration, and constant or recently declining mortality, a stable age distribution can be derived for the population and used as a standard for census evaluation purposes.

While no population is genuinely stable, this approach has proven to be quite useful in some developing country settings, particularly in cases where only one census has been conducted.

In countries which have experienced declining mortality over a significant number of years (a situation which exists in a number of countries), the direct application of stable population methods may not produce acceptable results. In such instances, the basic stable population approach can nevertheless be used if information on length of time during which mortality has been declining and the approximate rate of decline is available. This information can be used to derive adjustments to the implied stable age distribution to reflect the effects of declining mortality on the basis of quasi-stable population methods. The uses of stable and quasi-stable population methods for census evaluation purposes are described in section 5.2.

Finally, it is frequently the case that insight into broad patterns of error in census data can be obtained through the examination of the degree of internal consistency in the data. The basic premise behind this approach is that in order for a particular census item or characteristic to have been enumerated accurately, it should display plausible relationships with other characteristics. For example, the observation of higher population densities in rural than in urban areas and higher proportions of rural than urban households reporting having electric lighting will, in most developing countries, indicate coverage and/or content problems in the census.

Internal consistency may also be assessed through the examination of cross-tabulations of selected census characteristics. For example, one useful tabulation for assessing the plausibility of the recorded ages of young adults might be age by marital status. Under normal conditions, it would be expected that the proportion of persons who were "ever-married" would be very small at young ages and increase in a rather smooth progression over the age range in which marriage is culturally prescribed. Marked fluctuations and/or progression reversals across successive ages would be indicative of error in either the recorded ages or marital status. This approach may be applied in connection to any census characteristic whose relationship or association with age is fairly well known (for example, number of children ever born to females, duration of marriage, etc.). Applications of this approach to the evaluation of age reporting errors in sample surveys are described in Berggren et al. (1974), Ewbank (1981), and Goldman et al. (1979).

2.3 *Methods based upon comparison with a previous census or censuses*

A third class of demographic census evaluation techniques consists of methods in which the population "expected" at the time of a particular census is derived on the basis of the population enumerated in a previous census combined with information about one or more of the components of population change during the intercensal period. In the absence of independent data to assess directly the plausibility of census counts for one or more segments of the population, methods under this heading provide the primary basis for assessing census coverage error among the demographic methods described in this chapter. In view of the lack of data of sufficient quality to permit the direct evaluation of census results in many countries and the fact that many countries have now conducted at least their second census, one or more of these methods are likely to provide the firmest basis for comprehensive census evaluation in a sizable number of developing countries.

Evaluation methods based upon the analyses of successive censuses range from quite simple to relatively elaborate procedures. Perhaps the simplest procedure involves the assessment of the plausibility of the rate of population growth implied by the total population enumerated in two successive censuses. Given the total populations enumerated in two censuses (P_0 and P_1 respectively) taken t years apart, the implied rate of population growth, r, may be derived by solving the following expression for r:

(4.6) $$P_1 = P_0 \, exp \, (rt), \quad \text{or}$$
$$r = \frac{1}{t} \, \ell n \, \frac{P_1}{P_0}$$

This formula assumes an exponential rate of growth. Other mathematical formulae, described in U.S. Bureau of the Census (1975), can also be used.

The implied rate of growth is used to evaluate the degree of coverage in the censuses through the assessment of the plausibility of the estimate. An implausibly low implied rate of growth would be suggestive of net underenumeration in the second census relative to the first, which could have resulted either from underenumeration in the second census, overenumeration in the first census, or both. An implausibly high implied rate of intercensal population growth would be suggestive of the opposite pattern of relative or differential error in the two censuses.

If an independent estimate of the intercensal rate of population growth, r, is available (perhaps from vital statistics or a demographic survey conducted during the intercensal period), this estimate can be used in conjunction with equation (4.6) or similar mathematical function to derive an "expected" total population at the time of the second census against which to compare the population actually enumerated in the census. It should be noted that the comparison of the expected and enumerated census totals provides information on the mutual consistency of two census totals and the estimated rate of population growth, but not a direct estimate of the level of completeness of either of the individual censuses. This is because the first census, like the second census being evaluated, is subject to error. Accordingly, any observed discrepancy between the "expected" and the enumerated population in the second census suggests the presence of error in one or both censuses (or in the estimate of the rate of growth). If the expected population at the time of the second census is greater than the enumerated population, a net undercount of the second census <u>relative to</u> the first census is indicated. In this case, the implication would be that the second census was less completely enumerated than the first. A larger enumerated than expected population would suggest the opposite interpretation.

In countries in which vital registration systems are relatively complete or in which the degree of under-registration can be reliably estimated, the registered numbers of intercensal births, deaths, and international migrants (adjusted for under-registration, as required) may be used in conjunction with the population count from a previous census to derive an expected population at the time of a subsequent census. This approach makes use of the <u>population balancing equation</u>, which is essentially an accounting framework in which the number of births, B, and immigrants, I, are added to the population enumerated in the initial census, while the number of deaths, D, and emigrants, E, are subtracted from this population to derive the expected population at the time of the second census, as in equation (4.7).

$$(4.7) \qquad P_1 = P_0 + B - D + I - E \pm e$$

The residual term "e" in equation (4.7) is the quantity required to balance the equation. It provides an estimate of differential coverage error in the two censuses, although in actual practice errors in each of the components of the equation contribute to the residual error. See section 6.3 of chapter 5 for further information and an illustrative application of the balancing equation.

In cases in which usable vital registration data are unavailable but estimates of the prevailing levels of fertility and mortality can be derived from a demographic survey or from one or both censuses, estimated fertility and mortality rates may be used in conjunction with information on the size and composition of the population (by age and sex) at the time of a previous census to derive an expected population at the time of the second census. This approach, known as the cohort-component method, entails the "projection" of the first census population forward in time to the reference date of the second census on the basis of estimated schedules of fertility, mortality, and migration.

More specifically, the method entails the separate estimation of the effects of each of the components of population change on the size and composition of the population enumerated in the first census. The effects of mortality on the size of age and sex cohorts enumerated in the first census are estimated by applying life-table survival rates to each cohort to derive an expected count of survivors at the time of the second census. Births during the intercensal period are estimated by applying assumed age-specific fertility rates to the projected numbers of females in the reproductive years (ages 15 to 49) during the intercensal period and "survived" forward to the date of the second census on the basis of life-table survival probabilities. The effects of migration on the resulting size and distribution of the population are estimated by introducing actual data on the number of net migrants classified by age and sex or on the basis of an assumed schedule of age- and sex-specific net migration rates applied to the cohort populations enumerated in the first census.

The expected population derived in this manner is then compared with the population (by age and sex) enumerated in the second census. If it is possible to adjust the population enumerated in the first census for net census error prior to the application of this method, the comparison of the expected and enumerated populations at the time of the second census will yield estimates of net census error in the second census. If such adjustments are not possible, the method will yield estimates of relative or differential error in the two censuses.

Another approach which makes use of data from consecutive censuses involves the analysis of intercensal cohort survival rates. In the absence of census errors, the observed change in the size of each birth (age) cohort over the intercensal period should reflect the effects of mortality and migration alone. Accordingly, the observed ratio of the cohort population enumerated in the second census to that enumerated in the first census (the census cohort survival rate) should be consistent with that which would be considered "normal" under prevailing mortality and migration rates. In populations with negligible levels of net international migration, the observed intercensal cohort survival rates should approximate life-table survival probabilities for the intercensal period. In populations in which the level of net migration is significant, the "expected" survival probabilities must also take into account the volume and age-sex selectivity of migration. The degree to which the observed cohort survival rates deviate significantly from those expected on the basis of mortality and migration conditions is indicative of the degree of inconsistency between the two censuses.

Cohort survival rates may be used in a number of ways in census evaluation applications. The most direct application involves the assessment of the plausibility of the intercensal cohort survival rates implied by successive census enumerations. A variant of this approach entails the calculation of the ratio of the observed cohort survival rate to a life-table survival rate assumed to be representative of mortality conditions in the intercensal period. In the absence of census error and distortions due to migration, this ratio should equal 1.0. Significant deviations from this expected ratio which cannot be accounted for by migration signify the presence of error in one or both censuses. The uses of cohort survival techniques are described in further detail in section 6 of chapter 5.

The basic cohort survival rate approach has been extended to the estimation of age-reporting errors in two or three successive censuses (Demeny and Shorter, 1968; Das Gupta, 1975; Ntozi, 1978). Under this approach, information on the size and distribution of the population by age and sex in two (or three) successive censuses is used in conjunction with estimated life-table survival probabilities to derive a set of cohort correction factors which would make the two (or three) censuses and the assumed schedule(s) of survival rates jointly consistent. The correction factors derived are interpretable as estimates of net error in each of the censuses.

Finally, a method has been recently proposed (Preston and Hill, 1980; Luther, 1983) which makes use of data from two consecutive censuses and either death registration data or an estimated life-table for the intercensal period to estimate the relative level of completeness of coverage in the two censuses using ordinary least squares procedure. The method produces estimates of the correction factors for the two censuses which would be required to make the census counts and the recorded or estimated level of mortality during the intercensal period jointly consistent. The ratio of the correction factors derived in this fashion represents an estimate of differential completeness of coverage in the two censuses. This method is described in further detail in section 7 of chapter 5.

3. *TECHNIQUES FOR INDIRECT DEMOGRAPHIC ESTIMATION*

Several of the more powerful demographic techniques for census evaluation (those based upon analyses of successive censuses, in particular) require the availability of information on one or more demographic parameters in the population under study. Unfortunately, in many countries the statistical systems which traditionally provide such information (vital and immigration registration systems) are not yet developed to the point where the registration data are sufficiently accurate to be used directly.

The lack of usable direct data on fertility, mortality, and migration in many countries has led to the development of a set of statistical techniques which may be applied to incomplete and/or defective data of various types in order to provide "indirect" estimates of selected demographic parameters. The term indirect derives from the fact that such techniques produce estimates of demographic parameters on the basis of information which, in many cases, is only indirectly related to their value. Among these are methods which use demographic models and partial information about actual populations to derive estimates of other (unknown) parameters, techniques which provide adjustments for recorded data based upon consistency checks between selected data items, and techniques which use conventional data in unconventional ways (United Nations, 1983:2). In many countries, these techniques provide the primary basis for the estimation of demographic parameters and, as such, are likely to play an integral role in the census evaluation process.

This section summarizes how various types of census, survey, and registration data may be used in conjunction with one or more techniques of indirect estimation to derive estimates of basic demographic parameters. The purpose of presenting this material is to establish a link between different types of data, methods, and output (that is, estimates of fertility, mortality, and migration parameters). The reader is referred to United Nations (1983) and U.S. Bureau of the Census (1975) for more detailed descriptions (including assumptions and limitations of the methods described) and illustrative applications of these methods.

3.1 *Fertility estimation*

Techniques for the indirect estimation of fertility parameters may be categorized into three groups: (a) methods based upon census or survey information on the number of children ever born to women above the minimum age of childbearing, (b) stable population methods, and (c) methods based upon the "reverse-survival" of the population enumerated in a census.

Methods in the first group make use of responses to census or survey questions on the number of children ever born to women above the minimum age of childbearing (usually age 15) and an additional piece of information (which varies by method) to derive estimates of period and cumulative fertility rates. In the original procedure (Brass, 1964; see also Mortara, 1949) the reported cumulative fertility schedule in a census or survey is compared with a schedule of age-specific fertility estimated from census and/or survey questions on the number of live births during a specified period preceding the census or survey (usually 12 months) to derive adjusted age-specific and cumulative fertility schedules. The adjustments are derived based upon observation that certain features of the implied fertility schedules tend to be reported fairly accurately and may thus be used to adjust the remainder of the fertility schedules. In essence, the basic Brass procedure accepts as accurate the reported cumulative fertility of young women (ages 20 to 30 or 35) and uses this level of fertility to adjust the cumulative fertility level implied by the reported births in the previous period, which is assumed to represent fairly closely the actual shape of the period fertility schedule but not necessarily its level.

Subsequently, a number of alternative procedures have been developed using data for children ever born in an attempt to overcome some of the remaining weaknesses in the original method. Among these are (a) a variant based upon the comparison of cumulated "first-birth" fertility rates and the reported (in a census or survey) proportion of women in each age group reporting having borne at least one child; (b) a variant based on the comparison of the number of births registered by a cohort of women with the reported average parity of women in the same cohort in a survey or census; (c) a version based on the comparison of cohort fertility registered between two censuses with cohort parity increments between two censuses; (d) a method which measures period fertility based on cohort parity increments between two surveys (cf. Arriaga, 1983); and (e) a variant of the basic Brass age model based upon children ever born from survey or census data tabulated by the duration of marriage of the women reporting, which is thought to be reported more accurately than age in some settings.

Thus, depending upon the types of fertility information available and the extent to which certain assumptions regarding changes in fertility levels and types of errors in the data can be justified, the availability of basic survey or census information on the number of children ever born to women of reproductive age may be used in conjunction with several indirect estimation techniques to provide estimates of the level and age structure of fertility at a reference point close to a census date for use in census evaluation.

A second approach for the indirect estimation of fertility entails the use of stable population methods. As described in section 2, a population which has been exposed to approximately constant fertility and mortality and negligible net migration will, over a sufficiently long period of time, acquire a constant age distribution which is determined by the fertility and mortality rates to which it has been exposed. Accordingly, for the purposes of fertility estimation in a population which has been approximately stable, the age distribution from a current or previous census and an estimate of the rate of growth in the level of mortality in the population may be used to identify a model stable population which corresponds to two of the observed parameters in the population under study. (See United Nations, 1983, for a description of the various systems of model stable populations which are available for this purpose.) The level and age structure of fertility in the model stable population so identified may then be used as an estimate of the prevailing fertility regime in the population under study. In cases where mortality has been declining, quasi-stable population methods may be applied to derive appropriately adjusted estimates of fertility.

The third type of technique for indirect estimation of fertility consists of methods in which the number of children at any given age x enumerated in a census or survey, who represent the survivors of the births occurring x years prior to the census or survey, are "reverse-survived" on the basis of an assumed or estimated level of mortality to derive an estimate of the total number of births x years prior to the census or survey. This estimate is used as the numerator in any one of several fertility measures whose denominators are derived by reverse-surviving the total or some portion of the population (for example, females of childbearing age) to estimate the average population "at risk."

In the most direct application of the approach, the population enumerated in 5-year age groups is reverse-survived 5 or 10 years prior to the date of data collection on the basis of life-table survival probabilities. The resulting estimate of the number of births during each 5-year period prior to the census or survey is then used in conjunction with the estimated average population (either the total population or the number of women in each age group in the reproductive years) to derive estimates of crude birth rates.

A more detailed variant of this procedure involves the reverse-survival of the enumerated number of children and their mothers classified by single years of age back in time over a period of 10 or 15 years. Children are linked to their mothers on the basis of information on "relationship to head of household" typically included in census and survey household rosters. Conceptually, the "own-children" procedure is similar to that described above, but due to the more detailed nature of the required data, produces more detailed estimates (for example, age-specific fertility rates by single years of women). In actual practice, the own-children method is usually employed to derive estimates of age-specific fertility rates, with the basic reverse-survival procedure described above used to estimate crude birth rates (both of which provide useful information for census evaluation purposes).

For census evaluation purposes, fertility estimates derived on the basis of any of the approaches described above may be used in developing an expected population against which to compare an enumerated census population. Caution is advised, however, when using the results of the census under evaluation to derive estimates of fertility parameters. The reason for this is that the use of methods which are dependent upon aspects of the census which are to be evaluated will produce evaluation results which are biased to some extent. For example, if children aged 0 to 4 years are underenumerated in the census (which is quite frequently the case) and reverse-survival techniques are used to estimate the number of births during the 5 years prior to the census, the resulting fertility estimate will be biased downward. When this estimate is used to derive an expected count of persons aged 0 to 4 years at the time of the census, this expected population will be biased downward, as will the resulting estimate of census coverage error for this segment of the population.

Generally speaking, fertility estimates based upon children ever born data are preferable to those based upon reverse-survival methods for census evaluation purposes since the former tends to be less directly biased by errors in the census. While stable population methods are dependent upon the age distribution of the population recorded in the census (where this parameter is used to select a model stable population), the process of cumulation of the age distribution tends, at least partially, to dampen the biases in the estimates resulting from age misreporting or age-specific coverage errors. Similarly, while fertility estimates derived on the basis of analyses of census data for children ever born are clearly affected by errors in the census (coverage errors and age misreporting), the resulting estimates are usually less biased by these errors (unless the errors are correlated with fertility). Estimates from data for children ever born collected in a sample survey are even less dependent upon errors in the census and are generally less affected by non-sampling error, although subject to different types of errors (sampling error, for example).

The basic point is that for census evaluation purposes, the estimates of fertility levels used in developing an "expected" census population should be as independent of the census under evaluation as possible.

3.2 *Mortality estimation*

For descriptive purposes, techniques for the indirect estimation of mortality may be divided into three groups: (a) methods for the estimation of child mortality, (b) methods for the estimation of adult mortality, and (c) methods which produce estimates of the overall level of mortality (that is, both child and adult mortality). The methods in the first group, those producing estimates of infant/child mortality, derive from the method developed by Brass (1964) using census or survey information on numbers of children ever born by women of childbearing age (described in section 4.31) and numbers of children surviving. The basic procedure transforms the ratio of the number of surviving children to the total number of children ever born for women in each 5-year age category into a life-table measure of the probability of survival from birth to the average age of children in each 5-year age category of women. These estimates are adjusted, depending upon the fertility pattern, to yield estimates of the probability of dying from birth to the conventional ages 1, 2, 3, 5, 10, and 20. These values also provide a basis for estimating the reference point for each of the estimated $\ell(x)$ parameters assuming that mortality has been constant or has declined in a linear fashion (cf. Feeney, 1980).

Subsequently, a variety of modifications to and extensions of the basic method have been proposed. Sullivan (1972) and Trussell (1975) have developed alternative sets of adjustment multipliers to increase the flexibility of the method under differing fertility conditions. Extensions of the basic method include the development of a duration of marriage-based version of the Brass age model, the application of the basic method to data from two surveys to obtain an estimate of inter-survey child mortality, and a variant in which the fertility experiences of "true" cohorts are used to adjust the model-based adjustment multipliers used in the basic version of the method.

Recently, a method based upon data collected from women attending a medical/health center to bear a child regarding the survival of the previous child has been tested in several Latin American countries (see Pujol, 1985). This method uses information which can be quickly and inexpensively collected to provide an estimate of child mortality between the birth of the previous child and the date of birth of the newborn child (an average duration of about 30 months in high fertility societies).

Several approaches have been developed for the indirect estimation of adult mortality. One approach is based upon census or survey information on the orphanhood status of all persons interviewed (cf. Brass and Hill, 1973; Blacker, 1977). Under this approach, information on the proportion of respondents whose mother or father had died prior to the date of data collection is combined with information on the average age of childbearing women in the population to produce estimates of the probability of parental survival from the time of birth of respondents over a period of time equal to the respondent's age. These estimates are transformed into survival probabilities for conventional ages (for example, from ages 25 to 35, 40, 45, ..., 85). Because the method is based upon the mortality experience of the portion of the population who survived to bear children, the resulting estimates are conditional probabilities based upon survival to the mean age of childbearing. The method may be applied either to the estimation of female adult mortality (the maternal orphanhood method) based upon the survival status of the respondents mother or male adult mortality (the paternal orphanhood method) based upon similar information regarding the survival status of fathers.

A similar approach uses census or survey information on the survival status of first spouses to estimate adult survival probabilities (Hill, 1977). In the "widowhood" method, proportions of ever-married persons whose first spouse was still living at the time of data collection is used in conjunction with the information on mean age at first marriage (estimated by the singulate mean age at marriage - SMAM) to derive estimates of conditional survival probabilities for spouses from marriage to the date of the census or survey. These are then transformed into probabilities of survival from a reference age (age 20, for example, to ages 25, 30, 35, ..., 60) in much the same way as in the orphanhood method. Estimates for males are obtained based upon the reports of spousal survival status of female respondents, while estimates for females are derived on the basis of reports of male respondents.

Brass (1984) has recently shown that other estimates of adult mortality can be obtained from deductions regarding maternal survival even in the absence of direct information concerning survivorship. The method uses the excess of children recorded in a census over those reported by mothers as surviving to estimate the number of children whose mothers have died.

In general, the widowhood method is thought to produce more reliable estimates than the orphanhood method largely because it avoids some of the potentially serious biases associated with the orphanhood method (cf. United Nations, 1983:110).

An alternative approach for the estimation of adult mortality derives from the work of Brass (1975) and Preston et al. (1980). The two variants of this method are based upon the fact that in a stable population the rate of entry into the population aged x and above is equal to the rate of departure through death plus the stable population growth rate (which is the same for all values of x). The original Brass Growth Balance Method makes use of this relationship in developing a mathematical model from which the underreporting of deaths either in vital registration data or in survey reports on the number of deaths in the previous 12 months may be estimated by means of comparison with the age distribution of the population as recorded in a census. The basic method compares the age distribution of reported deaths with the age distribution of the population which is exposed to the underlying mortality regime to derive an estimate of the extent of underreporting of deaths. The resulting estimate may then be used to adjust the reported mortality data. The method proposed by Preston et al. attempts to increase the robustness of the Brass method in destabilized situations.

A third approach for the estimation of adult mortality uses the age and sex distributions from two successive census enumerations to derive estimates of the level of mortality during the intercensal period. The rationale for this approach is quite straightforward. In a population closed to migration in which two censuses have been taken i years apart, the population aged $x + i$ years at the time of the second census represents the "survivors" of the population aged x at the time of the first census. Accordingly, the probability of survival over the intercensal period at each initial age x may be estimated from the survival ratios implied by the two census age-sex distributions (assuming of course, accurate census enumerations).

The basic logic of this approach may be applied in a number of ways to estimate different mortality parameters. In the most direct application, implied cohort survival rates are compared with (smoothed by) model life table survival rates using the Coale-Demeny or other system of model life tables (see United Nations, 1983, for a description of the available systems of model life tables) to identify a level of mortality which is the most consistent with the implied census survival rates. The resulting "best" estimate of the mortality level may then be used to derive a full life table corresponding to the intercensal period.

When information on intercensal cohort deaths is available either from vital registration data or retrospective surveys, this information can be utilized along with population counts by age and sex from two successive censuses to derive estimates of both the extent of underreporting of deaths and differential coverage in the two censuses. The basic procedure entails the comparison of the observed census cohort survival rates between the two censuses with those expected on the basis of reported/registered deaths and the consistent adjustment of the census and mortality data. The application of this approach to census evaluation is described in greater detail in section 7 of chapter 5.

A third variant of the basic approach entails the use of the intercensal cohort growth rates implied by two successive censuses in a population that has been approximately stable to convert the average age distribution from the two censuses into the equivalent of a stationary population $({}_5L_x)$ life-table function from which various mortality parameters may be derived (Bennett and Horiuchi, 1981; Preston et al., 1980). While similar conceptually to the first variant described above, this variant uses the basic required data in an innovative manner to derive an estimate of expectation of life for the average intercensal population from which a full life table may be derived. This method recently has been extended to the case of destabilized populations given information on net migration rates (see Preston and Coale, 1982; Bennett and Horiuchi, 1984).

Finally, stable and quasi-stable population methods, as described above in connection with the estimation of fertility parameters, may be applied in the estimation of an overall level of mortality given the age-sex distribution from a census and an estimate of either the growth rate or the level of fertility in the population. In the case of mortality estimation, a model stable population would be selected on the basis of the recorded census age distribution and either the estimated growth rate or fertility level in the population and the mortality level implied by the two observed parameters accepted as an estimate of the overall level of mortality in the population under study. Under conditions of declining mortality, quasi-stable population adjustments may be applied if the length of time during and the rate at which mortality has been declining can be approximated.

Estimates of child and adult mortality derived on the basis of the methods described above may be combined into a consistent set of life-table survival probabilities following procedures described in United Nations (1983:147-155).

The caution noted in connection with fertility estimation regarding the need for the estimate(s) to be as independent as possible of the census under evaluation also applies to estimates of mortality. In the case of mortality estimation, estimates based upon the comparison of successive censuses are the least preferred because of their full reliance on the population count in the census being evaluated. While the estimates from other methods based upon the data from the census under evaluation also will be influenced by the quality of the census enumeration, the degree of dependency is generally lower.

3.3 *Migration estimation*

In the absence of usable statistics from border control systems, passenger counts, passport/visa application counts, etc., the estimation of international migration must depend upon census and survey data. Useful information for estimation purposes may be derived from responses to questions frequently included in censuses and household surveys, such as place (country) of birth, residence at a fixed prior date, duration of residence (in the country), period of immigration, and naturalization (citizenship) status. Information also may be derived from census or survey counts of nationals abroad and/or migration statistics of other countries.

Several approaches are available for utilizing census and survey information to estimate the volume of net migration during the period between two censuses or surveys. Separate estimates for immigration and emigration are generally not possible on the basis of census or survey data.

One basic approach makes use of the population balancing equation (see sections 5.3 and 6.3 of chapter 5) to provide an estimate of the volume of net migration. In the absence of census errors, the population enumerated at the time of a particular census will be equal to the population enumerated in a previous census plus the number of intercensal births minus the number of intercensal deaths plus or minus the number of intercensal migrants. Hence, if the numbers of intercensal births and deaths are known or can be estimated, these figures may be used in conjunction with the two census counts to estimate the number of net intercensal migrants as a residual in the balancing equation.

For census evaluation purposes, the estimates of migration obtained through this procedure will not be particularly useful because some portion of the residual estimate of migration will consist of differential coverage errors in the two censuses. As a result, the use of such migration estimates for census evaluation will produce biased estimates of net census coverage error.

Related methods based on data from two censuses may, however, prove useful for evaluation purposes. In countries where a large proportion of the net movement of population consists of movement of foreign-born persons, an estimate of net in-migration for an intercensal period may be obtained through the comparison of counts of foreign-born persons (ideally by age and sex) in successive censuses. Under this approach, the change in the size of the foreign-born population between two census dates which cannot be accounted for by mortality is used as an estimate of net migration between the two censuses. If the population of foreign-born persons enumerated in the second census exceeds that expected on the basis of the count enumerated in the first census and the prevailing level of mortality (introduced operationally in terms of life-table survival probabilities), net intercensal immigration would be inferred. A deficit of the enumerated foreign-born population in comparison with the expected count would imply net emigration during the intercensal period. Census counts of the population by citizenship may be used in a similar manner, but an adjustment for the depletion of the number of non-citizens due to naturalization during the intercensal period will be required.

In countries where the major portion of international migration consists of movement of nationals (native-born persons), the same approach may be applied to counts of nationals residing abroad in two censuses to derive an estimate of the net intercensal migration of nationals. In this case, a surplus of the enumerated population of nationals residing abroad in the second census over that expected on the basis of the count in the first census and the estimated level of intercensal mortality for this population would be indicative of net emigration from the country under study during the intercensal period, while a deficit of the enumerated to the expected population of nationals residing abroad would indicate net intercensal immigration.

In situations where there is significant movement of both nationals and foreign-born population, the separate estimates of net intercensal migration derived as described above may be combined into a single estimate for the intercensal period. The use of this approach assumes the appropriate questions have been included in both censuses.

It should be noted that these estimates are subject to the same limitations noted above in connection with the population balancing equation, namely that the resulting estimates of migration will be affected to some extent by differential levels of coverage and content error in the two censuses. However, since the numbers of persons involved in the estimation procedure are substantially smaller than the residual method based upon the entire population, the resulting bias will often have a proportionally small effect on the resulting estimates of census error.

Census or survey questions on place of residence at a fixed prior date, for example 5 or 10 years prior to the census or survey, and period of immigration may be used in a more direct fashion to estimate the level of immigration in specified periods preceding the census or survey. However, these data suffer from several methodological difficulties. First and foremost among these is that the respondents interviewed represent the "survivors" of the stream of in-migrants and not the actual numbers. Immigrants who had died or had emigrated prior to the date of data collection are missed in these data, and thus bias the resulting estimate of immigration. While the mortality bias may be minimized by "reverse-surviving" the enumerated population of immigrants over appropriate periods depending upon the reference periods used in the census or survey questions, the problem of subsequent emigration prior to the date of data collection remains. Further, these types of data provide no basis for measuring the movement of nationals. In summary, while such data provides some information on migration in various periods prior to a census or survey, it should be used with great caution for census evaluation purposes.

In some cases, usable data on emigration from a country may be derived from immigration statistics or censuses or surveys of other countries. Where the data may be assumed to be reasonably accurate, statistics on numbers of immigrants and emigrants of another country may be used directly as a substitute for such data in the country under study. Due to the generally low quality of such data, however, adjustments for suspected biases are usually required. Data from one or two censuses or surveys in another country may be used in the manner described above, assuming that the data are published or otherwise available in a form that indicates the flow of population between the country of interest and the country for which the data are available.

Several countries have attempted to measure emigration of nationals by including questions on relatives residing abroad in censuses and surveys. To date, however, the results of this approach have been for the most part disappointing. The major methodological problems encountered under this approach involve duplicate reporting of individuals migrating from households in which one or more persons remain behind and undercoverage of cases in which the entire household has migrated.

Finally, migration researchers have recently begun to explore the possibilities of using network (multiplicity) sampling techniques in survey efforts to estimate the volume of migration within specified periods (Goldstein and Goldstein, 1981). Under this approach, respondents who are "captured" in the survey sample are asked to report on the movement of persons with whom they are acquainted (cf. Sirken, 1970 and 1972). While there are several methodological problems inherent in this approach which must be minimized to the extent feasible, the approach has considerable potential, particularly in cases where alternative sources of information are not available.

4. REFERENCES

Anderson, B., and Silver, B. 1985. "Estimating Census Undercount from School Enrollment Data: An Application to the Soviet Censuses of 1959 and 1970." *Demography* 22(2):280-288.

Bennett, N., and Horiuchi, S. 1984. "Mortality Estimation from Registered Deaths in Less Developed Countries." *Demography* 21(2):217-233.

Berggren, G.; Murthy, N.; and Williams, S. 1974. "Rural Haitian Women: An Analysis of Fertility Rates." *Social Biology* 21:332-339.

Bhat, P.N.; Preston, S.; and Dyson, T. 1984. *Vital Rates in India, 1961-1981.* Washington, D.C.: National Academy of Sciences, Committee on Population and Demography.

Blacker, J.G.C. 1977. "The Estimation of Adult Mortality in Africa from Data on Orphanhood." *Population Studies* 31(1):107-128.

Brass, W. 1964. "Uses of Census or Survey Data for the Estimation of Vital Rates." (E/CN.14/CAS.4/057). Paper prepared for the African Seminar on Vital Statistics, Addis Ababa.

-----. 1975. *Methods for Estimating Fertility and Mortality from Limited and Defective Data.* Chapel Hill: Carolina Population Center and Laboratories for Population Studies.

-----. 1984. "Mortality in China Over the Past 50 Years: Indirect Estimates from the 1982 Census." Princeton University, Office of Population. Mimeo.

Brass, W., and Hill, K.H. 1973. "Estimating Adult Mortality from Orphanhood." In Vol. 3 of the *Proceedings of the International Population Conference*, 111-123. Liege: International Union for the Scientific Study of Population.

Coale, A.J., and Demeny, P. 1966. *Regional Model Life Tables and Stable Populations.* Princeton: Princeton University Press.

Das Gupta, P. 1975. "A General Method for Correction for Age Misreporting in Census Populations." *Demography* 12(2):303-312.

Demeny, P., and Shorter, F. 1968. *Estimating Turkish Mortality, Fertility and Age Structure.* Istanbul: Istanbul University.

Ewbank, D. 1981. *Age Misreporting and Age-Selective Underenumeration: Sources, Patterns, and Consequences for Demographic Analyses.* Washington, D.C.: National Academy of Sciences, Committee on Population and Demography.

Feeney, G. 1980. "Estimating Infant and Childhood Mortality under Conditions of Changing Mortality." *Population Studies* 34(1):109-128.

Goldman, N.; Coale, A.J.; and Weinstein, M. 1979. *The Quality of Data in the Nepal Fertility Survey*. Scientific Report No. 6. London: World Fertility Survey.

Goldstein, S., and Goldstein, A. 1981. "The Use of the Multiplicity Survey to Identify Migrants." *Demography* 18(1):67-84.

Hill, K. 1977. "Estimating Adult Mortality Levels from Information on Widowhood." *Population Studies* 31(1):75-84.

-----. 1985. "Overview: New Methods of Adjustment and New Survey Approaches to Demographic Estimation." In Vol. 4 of the *Proceedings of the International Population Conference*, 5-9. Florence: International Union for the Scientific Study of Population.

Luther, N. 1983. "Measuring Changes in Census Coverage in Asia." *Asian and Pacific Census Forum* 9(3):1-16.

Ntozi, J. 1978. "The Demeny-Shorter and Three-Census Methods for Correcting Age Data." *Demography* 15(4):509-522.

Preston, S.H., and Coale, A.J. 1982. "Age Structure, Growth, Attrition, and Accession: A New Synthesis." *Population Index* 48(2):217-259.

Preston, S.; Coale, A.J.; Trussell, T.J.; and Weinstein, M. 1980. "Estimating the Completeness of Reporting of Adult Deaths in Populations that are Approximately Stable." *Population Studies* 46(2):179-202.

Preston, S., and Hill, K. 1980. "Estimating the Completeness of Death Registration." *Population Studies* 34:349-366.

Pujol, T.M. 1985. "New Methods of Evaluating and Adjusting Conventional Demographic Data." In Vol. 4 of the *Proceedings of the International Population Conference*, 11-22. Florence: International Union for the Scientific Study of Population.

Sirken, M. 1971. "Household Survey with Multiplicity." *Journal of the American Statistical Association* 65:257-266.

-----. 1972. "Stratified Sample Surveys with Multiplicity." *Journal of the American Statistical Association* 67:224-27.

Sullivan, J. 1972. "Models for the Estimation of the Probability of Dying between Birth and Exact Ages of Early Childhood." *Population Studies* 26(1):79-97.

Trussell, T.J. 1975. "A Re-estimation of the Multiplying Factors for the Brass Technique for Determining Childhood Survivorship Rates." *Population Studies* 29(1):97-108.

United Nations. 1967. *Methods of Estimating Basic Demographic Measures from Incomplete Data*. Manual 4. New York: United Nations.

-----, Population Division; and the U.S. National Academy of Sciences. 1983. *Indirect Techniques for Demographic Estimation*. Manual 10. New York: United Nations Population Division and the U.S. National Academy of Sciences.

U.S. Bureau of the Census. 1975. *The Methods and Materials of Demography*. Washington, D.C.: U.S. Bureau of the Census.

-----. 1983. *Estimating Fertility from Children Ever Born by Age of Mother*. Washington, D.C.: U.S. Bureau of the Census.

Chapter 5. APPLICATION OF SELECTED DEMOGRAPHIC TECHNIQUES FOR CENSUS EVALUATION

1. *INTRODUCTION*

This chapter describes and provides illustrative example applications of some of the more useful demographic techniques for census evaluation in developing country settings. The chapter begins with a description of basic graphical techniques for the visual inspection of census age and sex distributions. The application of stable population theory to the evaluation of census age and sex distributions is described next. Following this discussion, four evaluation approaches based upon the comparison of data from successive censuses are illustrated. The first of these uses registered numbers of births, deaths, and net migrants for the period between the two censuses to derive an "expected" population at the time of the second census which is used as a "standard" for evaluation purposes. The second method is based upon the use of estimates of fertility, mortality, and migration to project the population enumerated in the first census to the date of the second census. This projected population is used as a standard against which to evaluate the second census. The third approach involves the assessment of the plausibility of the two census age distributions through the examination of implied intercensal cohort survival rates. The final approach considered uses regression procedures to estimate the coverage "correction" factors for the two censuses which would be required to make the two census counts (by age) mutually consistent with each other and the underlying level of mortality in the intercensal period. The ratio of these correction factors represents an estimate of relative completeness of coverage in the two censuses.

With the exception of the first approach considered (graphical analyses), the presentation of each method is arranged in the following fashion: (a) a discussion of the basis or rationale of the method; (b) the data required for its application; (c) the step-by-step computational procedures involved; (d) one or more examples of the application of the method; and (e) a discussion of the uses and limitations of the method.

Because of the difficulties involved in obtaining separate estimates of coverage and content error using these techniques, no attempt was made to organize the chapter around the measurement objectives of the methods described. Instead, the utility of each method for measuring coverage and content error is indicated in the discussion of the method and, where possible, illustrated in the example application of the method.

2. *ANALYSES OF AGE-SEX DISTRIBUTIONS*

Evaluation of the "reasonableness" of the distribution of the population enumerated in a census by age and sex can provide considerable insight into the quality of the census enumeration. The reason for this is that the age and sex distribution of a population, which is determined by the levels of fertility, mortality, and international migration to which the population has been exposed, follows a well-known and fairly predictable pattern (in the absence of error in the data). Accordingly, significant discrepancies in a census age-sex distribution which cannot be accounted for by extraordinary events (wars, famines, etc.) or other distorting factors (significant levels of age- or sex-selective international migration or abrupt changes in fertility/mortality levels, for example) are usually indicative of census errors.

Detailed evaluation of age-sex distributions can be undertaken in the absence of any other information. However, an important limitation of such analyses is that it is often difficult to determine the sources of observed discrepancies in the data. Nevertheless, because errors in census age-sex data tend to follow fairly common patterns whose sources are known from previous research, it is often possible to conclude with a fairly high degree of confidence that either census errors or other distorting factors were primarily responsible for observed distortions in various segments of the distributions (although it is less frequently the case that direct numerical estimates of coverage and content error rates can be derived).

In this section, several approaches for assessing the quality of census age and sex data are described. First, the uses of age-sex pyramids and related graphical techniques to visually inspect the data are described. Second, a somewhat more quantitative approach based upon the analysis of age- and sex-ratios is illustrated. The uses of selected indices of overall accuracy and of the magnitude of particular patterns of error, "age heaping," are then summarized. The section concludes with a discussion of the practical uses and limitations of these methods.

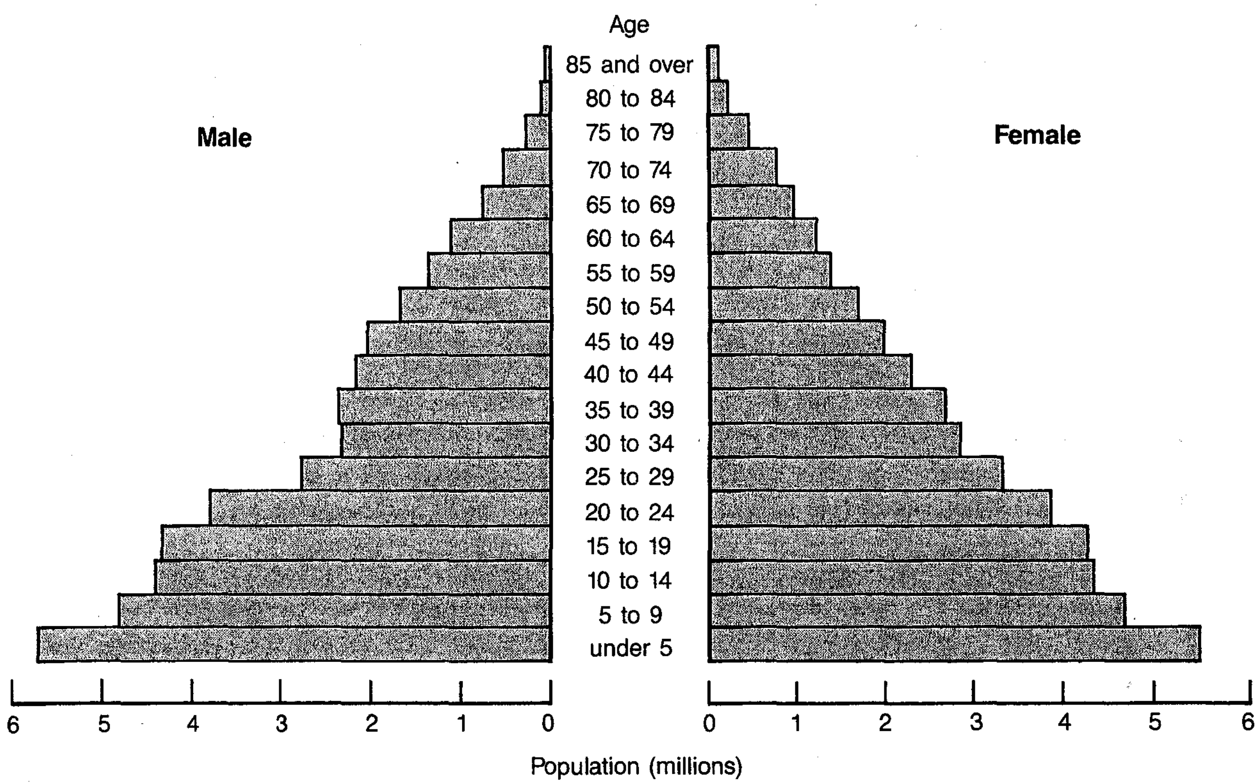
2.1 *Graphical analysis*

One of the more basic procedures for assessing the quality of census data on age and sex is to visually examine the data using graphical techniques. Perhaps the most widely-used among these is the use of the population pyramid, which displays the size of the population enumerated in each age group (or cohort) by sex. In a population which has experienced relatively constant levels of fertility and mortality and at most modest levels of migration, this display takes the form of a pyramid. The wideness of the base of the pyramid is determined by the level of fertility in the population, while the rapidity at which the pyramid converges to its peak is a function of previous levels of mortality and fertility.

The population distribution enumerated in the 1950 Census of Japan (figure 5-1) illustrates the typical shape of a population age pyramid (shown in 5-year age groups) in a population with relatively high levels of fertility and mortality and negligible migration and in which the census enumeration is of relatively high quality. The major distortion in the pyramid for Japan, the deficit of males aged 25 to 39, is the result of war casualties.

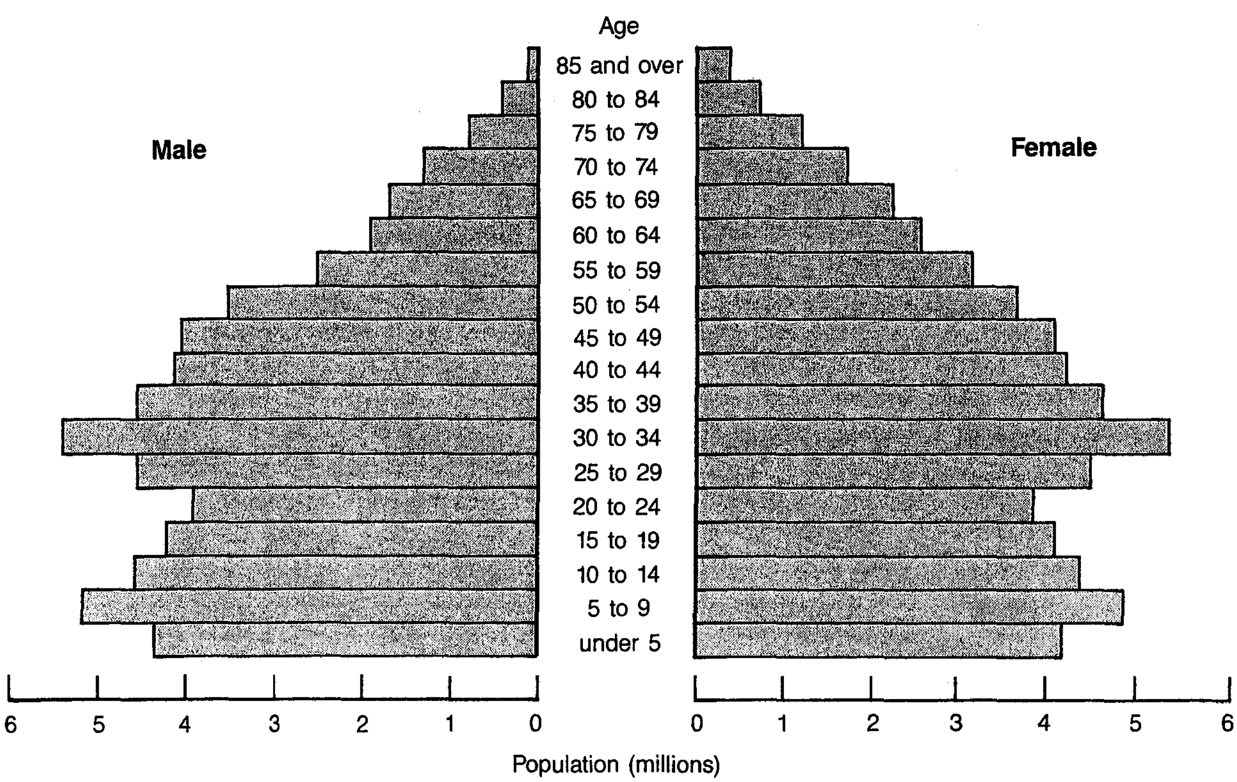
The effects of declining fertility and mortality on the age distribution may be observed by comparing figure 5-1 with figure 5-2, which shows the age distribution of the 1980 census population of Japan. The narrower base of the pyramid in the 1980 census reflects the rapid decline in fertility which took place during the last 30 year period. The irregular sizes of the cohorts under the age of 15 are largely the result of distortions among older cohorts in the Japanese population caused initially by war, known as an "echo" effect. Over the course of several generations these irregularities will be diminished and a smooth pyramid with a narrower base than that shown in figure 5-1 will eventually evolve (assuming, of course, that no

Figure 5-1. CENSUS POPULATION OF JAPAN BY AGE AND SEX, 1950



Source: Japan Statistics Bureau, 1982, 1980 Population Census of Japan, Volume 3, Results of the First Basic Complete Tabulation, Part 1, Whole Japan, Table 3.

Figure 5-2. CENSUS POPULATION OF JAPAN BY AGE AND SEX, 1980



Source: Japan Statistics Bureau, 1982, 1980 Population Census of Japan, Volume 2, Results of the First Basic Complete Tabulation, Part 1, Whole Japan, Table 2.

extraordinary event or further major swings in fertility or mortality levels occur).

In contrast with the case of Japan, where the irregularities in the population pyramid can be attributed largely to declining fertility levels and periods of high mortality associated with war, the causes of the irregularities in the 1975 de facto census population of Yemen (figure 5-3) are less clear. While some portion of the deficit of population (particularly among males) ages 15 to 35 is likely attributable to labor migration to other middle eastern countries, some of the other distortions in the distribution are likely the result of census error. The deficit of children under 5 years of age, for example, suggest that young children may have been underenumerated and/or transferred into the 5 to 9 age category through age reporting errors. However, the possibility that the deficit of young children reflects the relatively small size of the cohorts in the peak fertility years (ages 20 to 24) cannot be dismissed solely on the basis of figure 5-3.

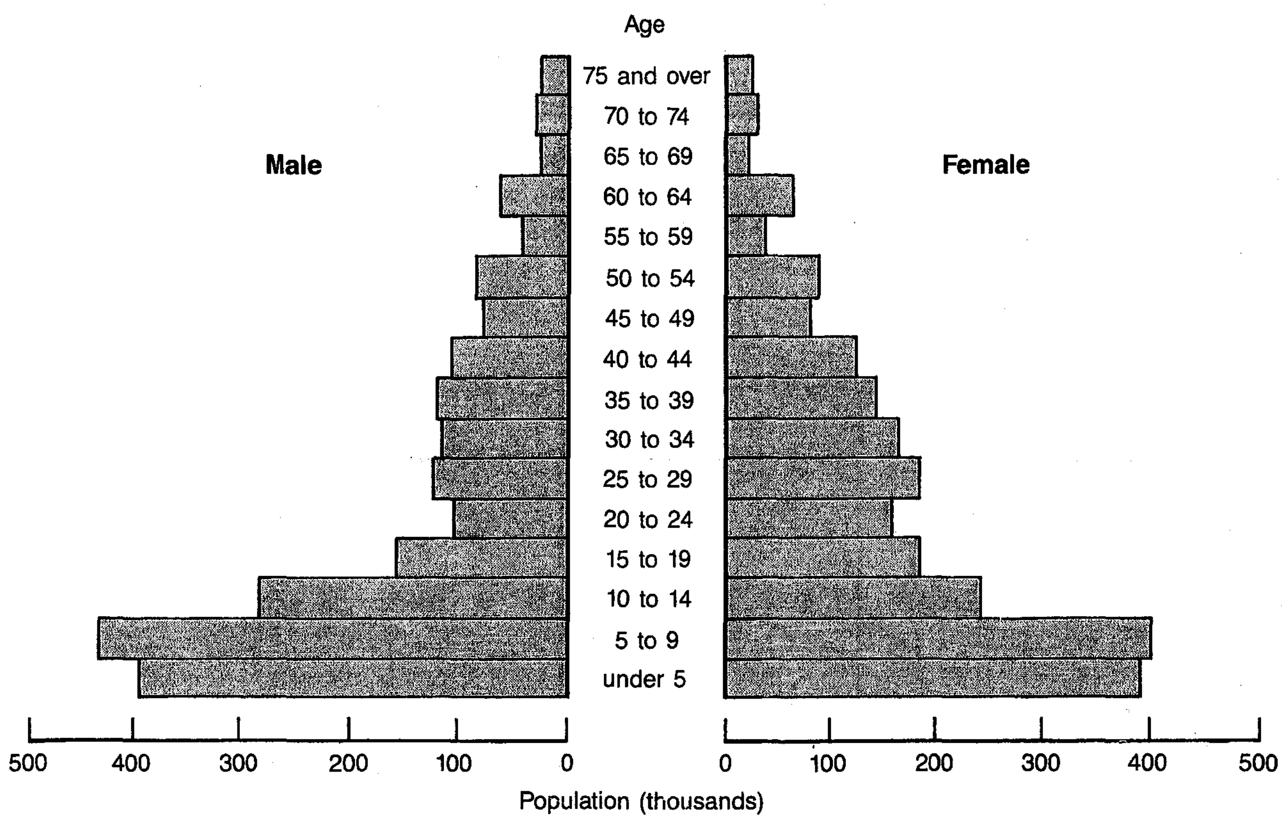
In actual practice, additional investigation would be required to assess whether the size of the enumerated cohort of young children is consistent with the size of the enumerated population in the childbearing years (ages 15 to 49) and estimates of the current levels of fertility and infant mortality in Yemen. (See section 5.4 for further discussion of this procedure.) Insight into this issue also can be gained through the examination of the census data tabulated by single years of age (presented later in this section).

While the interpretation of the small size of the cohort of children under 5 years of age is unclear solely on the basis of figure 5-3, the irregular sizes of both male and female cohorts above age 45 is very likely to indicate the presence of error in the data. Because the pattern of alternating excesses and deficits among successively older cohorts, age misreporting (content error) would seem likely to be the major source of error in this segment of the population.

The age-sex data from the 1975 census of Yemen will be used throughout this section to illustrate the various techniques discussed. For comparative purposes, data from the 1980 census of Argentina (see figure 5-4) also will be used. As indicated in figure 5-4, the age pyramid from the Argentina census conforms more closely to a smooth distribution, although some irregularities are evident. The actual data used to develop these pyramids are presented in figures 5-5 and 5-6 to permit the reader to follow the computational procedures involved for the graphical and numerical techniques covered in this section.

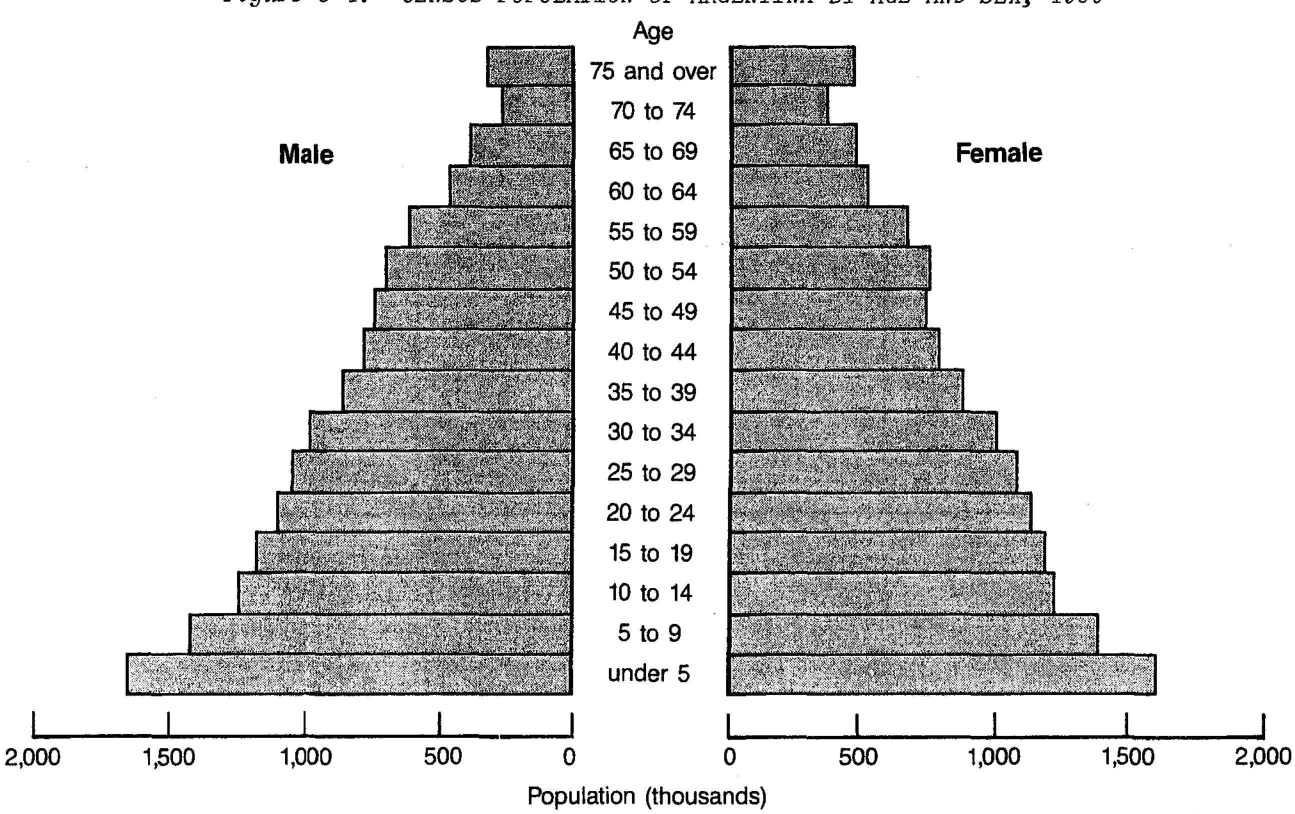
An examination of the data by single years of age provides a closer look at the patterns of error suggested by the age pyramids based upon 5-year grouped data (see figures 5-7 and 5-8). The Yemen data shown in figure 5-7 suggest rather strongly, as was suggested by figure 5-3, that young children (under age 2) were underenumerated in the census and that the magnitude of age reporting errors among adults is significant. The particular form of age error illustrated in figure 5-7, a tendency for respondents to report ages ending in particular digits (often 0 and 5, as in the case of Yemen), is know as <u>digit preference</u> or <u>age heaping</u> and is usually observed to some degree in most censuses. A less pronounced degree of heaping on ages ending in the digits 0 and 5 also is apparent in the Argentine census (figure 5-8). Otherwise, the recorded age distribution of the population enumerated in the Argentine census reveals only minor irregularities. The major uncertainty here might be the somewhat smaller size of the 5 to 9 and 10 to 14 age ranges in comparison

Figure 5-3. DE FACTO CENSUS POPULATION OF YEMEN BY AGE AND SEX, 1975



Source: "Evaluation and Analysis of the 1975 Population and Housing Census: Testing the Accuracy of Age-Sex Statistics," Yemen Arab Republic, Central Planning Organization, Statistics Dept., 1980, Table 1.

Figure 5-4. CENSUS POPULATION OF ARGENTINA BY AGE AND SEX, 1980



Source: Censo Nacional de Poblacion y Vivienda 1980, Serie D, Resumen Nacional, Table G.2.

Figure 5-5. DE FACTO POPULATION OF YEMEN BY AGE AND SEX, 1975

Age	Males	Females	Age	Males	Females	Age	Males	Females
0 to 4	392,947	393,969	30 to 34	113,706	165,673	60 to 64	59,834	65,237
0	69,778	67,688	30	84,948	129,389	60	51,266	58,138
1	63,320	64,208	31	4,118	5,087	61	1,616	1,350
2	85,582	87,505	32	12,915	16,240	62	3,190	2,754
3	85,788	87,674	33	5,858	7,179	63	1,949	1,632
4	88,479	86,894	34	5,867	7,778	64	1,813	1,363
5 to 9	432,284	401,496	35 to 39	117,589	143,951	65 to 69	24,396	22,746
5	92,529	87,951	35	86,235	108,643	65	19,575	18,819
6	86,834	82,161	36	7,160	7,660	66	1,113	780
7	93,404	86,736	37	7,400	7,180	67	1,292	911
8	92,529	86,697	38	12,141	14,115	68	1,675	1,415
9	66,988	57,951	39	4,653	6,353	69	741	821
10 to 14	281,036	243,576	40 to 44	104,442	125,601	70 to 74	28,641	32,041
10	78,723	69,736	40	84,430	105,820	70	24,811	28,935
11	27,443	23,279	41	3,474	3,390	71	758	635
12	88,956	73,391	42	8,657	8,200	72	1,468	1,309
13	46,681	39,782	43	4,956	5,166	73	875	639
14	39,233	37,388	44	2,925	3,025	74	729	523
15 to 19	153,427	185,793	45 to 49	75,741	81,976	75 to 79	10,336	9,781
15	64,684	68,858	45	58,545	64,352	75	8,422	8,221
16	27,819	35,275	46	3,822	3,633	76	591	464
17	16,513	20,802	47	3,726	3,341	77	430	268
18	34,393	48,738	48	6,646	6,896	78	592	486
19	10,018	12,120	49	3,002	3,754	79	301	342
20 to 24	101,488	159,447	50 to 54	82,106	91,371	80 to 84	11,509	14,215
20	56,815	96,461	50	68,461	78,989	80	10,305	13,131
21	6,791	8,453	51	2,292	2,087	81	283	277
22	17,733	24,424	52	5,254	4,589	82	452	428
23	10,724	15,286	53	3,019	2,627	83	235	197
24	9,425	14,823	54	3,080	3,079	84	234	182
25 to 29	120,335	185,731	55 to 59	39,915	39,211	85 to 89	2,802	2,822
25	73,577	121,819	55	29,753	29,873	85	2,315	2,329
26	10,008	13,012	56	2,971	2,398	86	148	143
27	12,097	14,383	57	2,236	1,827	87	126	131
28	19,539	28,237	58	3,469	3,369	88	108	108
29	5,114	8,280	59	1,486	1,744	89	105	111

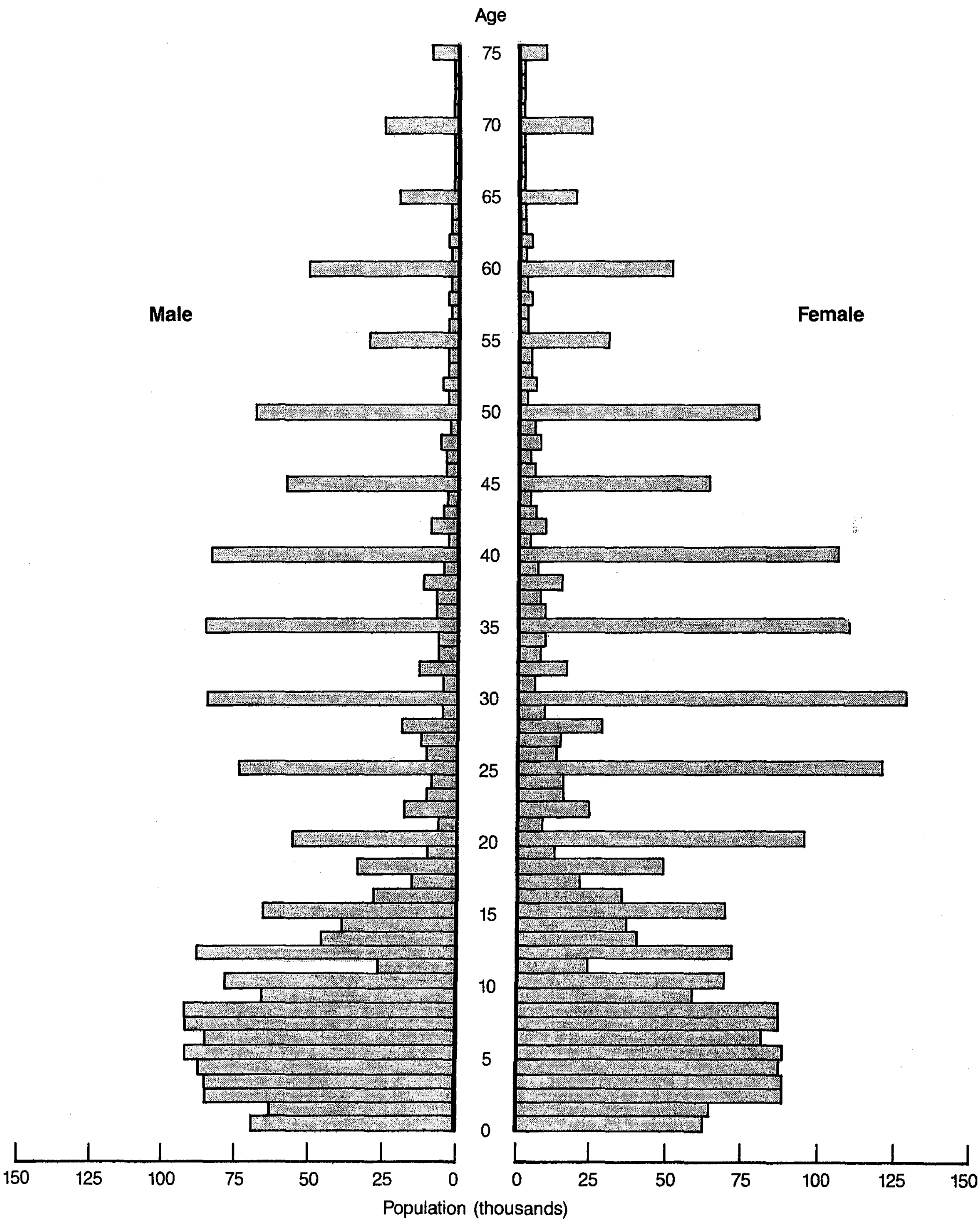
Source: Yemen Arab Republic (1980), table 1.

Figure 5-6. CENSUS POPULATION OF ARGENTINA BY AGE AND SEX, 1980

Age	Males	Females	Age	Males	Females	Age	Males	Females
0 to 4	1,640,005	1,600,836	30 to 34	979,816	994,377	60 to 64	469,993	532,126
0	342,896	334,747	30	217,870	225,299	60	116,525	136,880
1	321,274	312,438	31	189,095	188,118	61	87,360	91,888
2	331,097	324,020	32	196,979	202,289	62	91,919	103,578
3	333,115	324,913	33	195,215	195,518	63	87,836	100,047
4	311,623	304,718	34	180,657	183,153	64	86,353	99,733
5 to 9	1,407,213	1,376,543	35 to 39	855,293	869,013	65 to 69	397,567	476,239
5	301,738	295,023	35	186,968	192,803	65	92,407	111,053
6	279,854	274,443	36	174,601	177,361	66	81,938	96,304
7	276,410	269,268	37	164,553	167,640	67	78,978	93,040
8	278,890	271,300	38	168,668	171,788	68	75,229	92,660
9	270,321	266,509	39	160,503	159,421	69	69,015	83,182
10 to 14	1,240,209	1,215,962	40 to 44	772,913	775,702	70 to 74	279,279	355,241
10	267,240	262,024	40	178,937	183,129	70	70,483	94,098
11	250,290	248,446	41	143,126	140,992	71	55,709	65,087
12	247,664	241,187	42	157,569	160,115	72	57,168	71,439
13	236,129	233,656	43	149,899	149,695	73	49,591	64,779
14	238,886	230,649	44	143,382	141,771	74	46,328	59,838
15 to 19	1,173,841	1,167,647	45 to 49	748,046	748,741	75 to 79	181,291	245,572
15	240,668	239,411	45	154,777	157,083	75	47,921	63,152
16	238,698	234,287	46	141,947	141,240	76	40,815	53,432
17	234,614	234,996	47	142,978	143,109	77	34,813	44,103
18	241,064	231,823	48	153,776	155,879	78	30,761	45,457
19	218,797	227,130	49	154,568	151,430	79	26,981	39,428
20 to 24	1,099,810	1,124,347	50 to 54	709,825	749,808	80 to 84	85,524	138,377
20	218,182	228,487	50	158,858	175,486	80	26,864	44,508
21	214,393	218,027	51	133,443	134,098	81	17,073	25,319
22	221,197	225,674	52	145,612	153,652	82	15,595	25,504
23	227,106	228,426	53	138,249	145,204	83	13,646	22,102
24	218,932	223,733	54	133,663	141,368	84	12,346	20,944
25 to 29	1,050,065	1,074,218	55 to 59	620,972	659,561	Over 84	44,321	87,153
25	215,544	223,924	55	133,447	143,792			
26	204,093	212,604	56	129,616	137,016			
27	211,619	215,975	57	122,653	126,957			
28	212,382	214,317	58	121,258	131,210			
29	206,427	207,398	59	113,998	120,586			

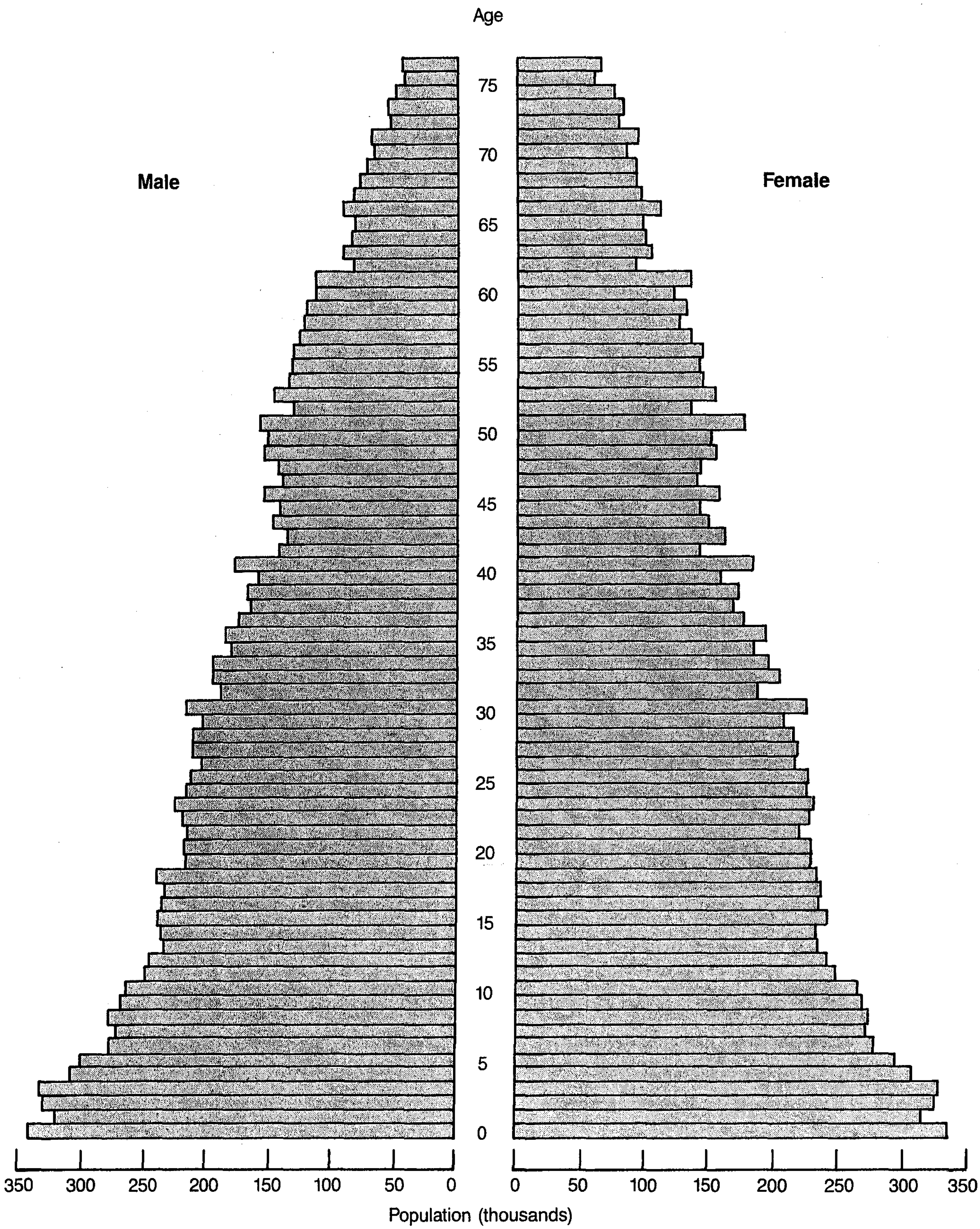
Source: Argentina (1982), table G.2.

Figure 5-7. DE FACTO CENSUS POPULATION OF YEMEN BY SEX AND SINGLE YEARS OF AGE, 1975



Source: Yemen Arab Republic (1980), table 1.

Figure 5-8. CENSUS POPULATION OF ARGENTINA BY SEX AND SINGLE YEARS OF AGE, 1980



Source: Argentina (1982), table G.2

with the 0 to 4 and 15 to 19 ranges. The relatively large size of the population aged 0 to 4 years in comparison with those aged 5 to 9 years might be due to a recent increase in fertility which has been noted in at least one other Latin American country (Chile). The modest size of the population aged 10 to 14 years relative to adjacent cohorts suggests the presence of error in the census, or immigration in the 15 to 19 age category (possibly from Bolivia or Paraguay), although the cause is not apparent from the age pyramid alone.

An alternative to the population pyramid is the graphical display of cohort sizes by age shown for Yemen and Argentina in figures 5-9 and 5-10 (shown in single years of age). This type of graphical display serves essentially the same purpose as the population or age pyramid, but may be somewhat easier to analyze particularly when two or more census age distributions are being compared in the same graph.

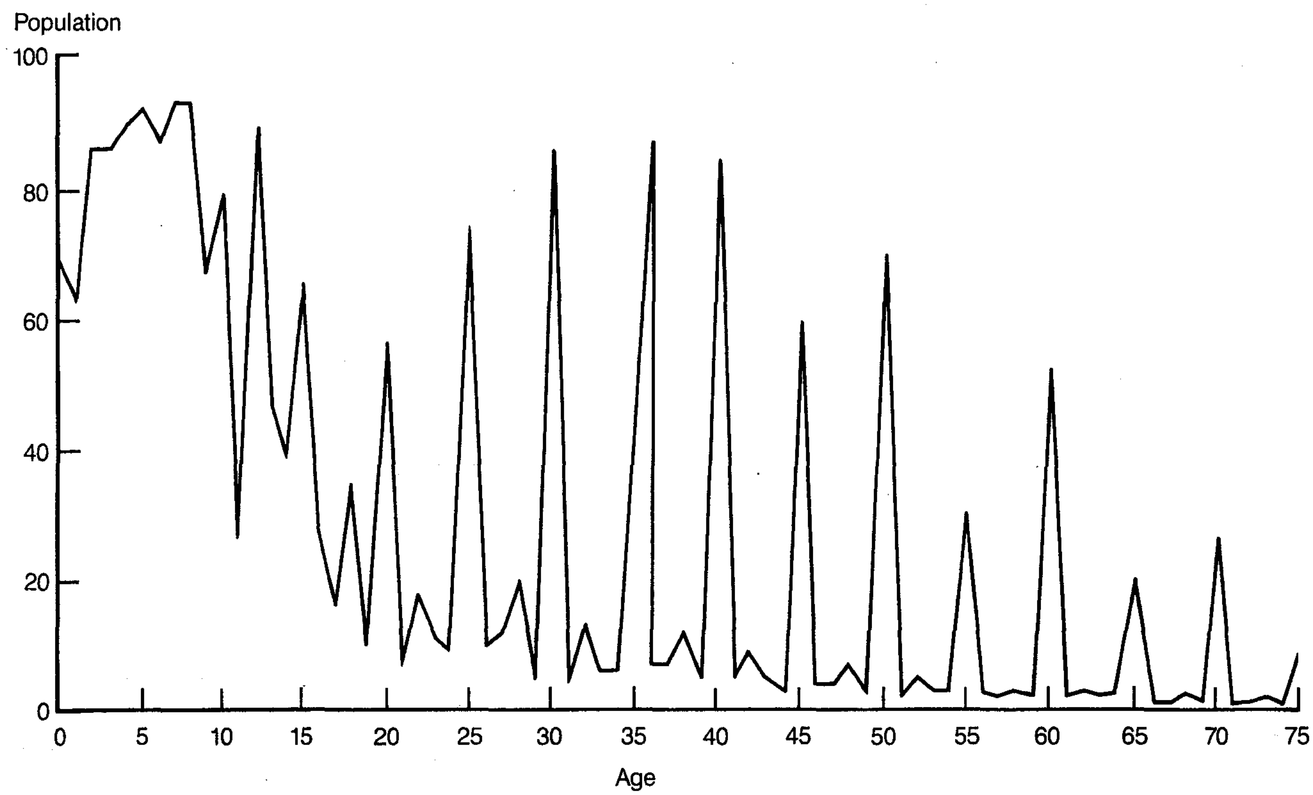
When two or more censuses are available, the graphical technique illustrated in figures 5-9 and 5-10 may be extended to examine the consistency of the age distributions in successive censuses. This procedure, known as graphical cohort analysis, entails the plotting of the size of actual cohorts in each of the censuses being compared on "semi-log" paper (the y-axis of the graph is expressed on a logarithmic scale). Due to mortality, the size of each cohort should decline in successive censuses and, in the absence of errors in the censuses and migration, the lines for the successive censuses should follow the same pattern. An important advantage of this approach is that because actual cohorts are followed over time, it is generally possible to discern the effects of extraordinary events and other distorting factors on the census age-sex distributions.

A graphical cohort analysis is illustrated in figure 5-11 using population counts of males by age (in 5-year age groups) from the 1970, 1975, and 1980 censuses of Turkey. The points on the line for each census are population counts for each cohort plotted at the mid-point of the 5-year interval during which the cohort was born. For the 0 to 4 age group in the 1980 census, for example, the census count of 3.05 million is plotted at the midpoint of the 1975-1980 period during which they were born. This procedure is repeated for each cohort in the 1980 census and then for the 1975 and 1970 censuses.

In the absence of census errors or migration, the three lines should follow the same trend and not cross each other, with the line corresponding to the 1970 census on top and the line for the 1980 census on the bottom. In fact, the lines in figure 5-11 depart somewhat from this expected pattern, most notably at the youngest and oldest ages. The fact that the lines for the 1975 and 1970 census begin lower than and cross over the line for the 1980 census between the ages of 5 to 9 is suggestive of underenumeration of children aged 0 to 4 years and/or a systematic transfer of infants and young children into the 5 to 9 age category. More extreme volatility is observed at the older ages, which is typically associated with proportionately larger errors made in reporting ages among older populations.

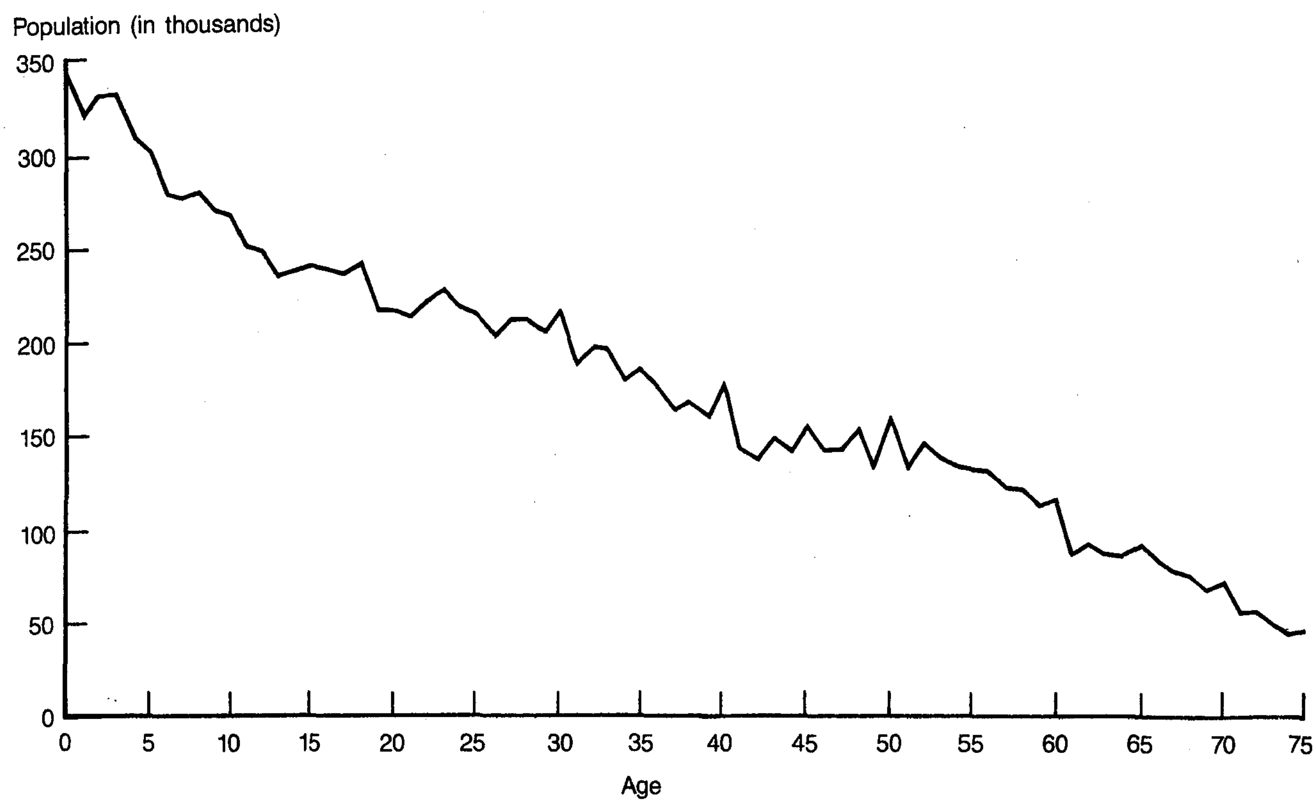
One of the more useful features of graphical cohort analysis is that the comparison of successive censuses provides the analyst with at least a partial basis for separating historical artifacts in the age distribution from the effects of census errors. In the case of Turkey, for example, the fact that smaller surviving cohorts are observed in all three censuses for persons born during war years (1913-1917 and 1933-1943) suggests that these distortions largely reflect

Figure 5-9. DE FACTO CENSUS POPULATION OF YEMEN, MALES BY SINGLE YEARS OF AGE, 1975



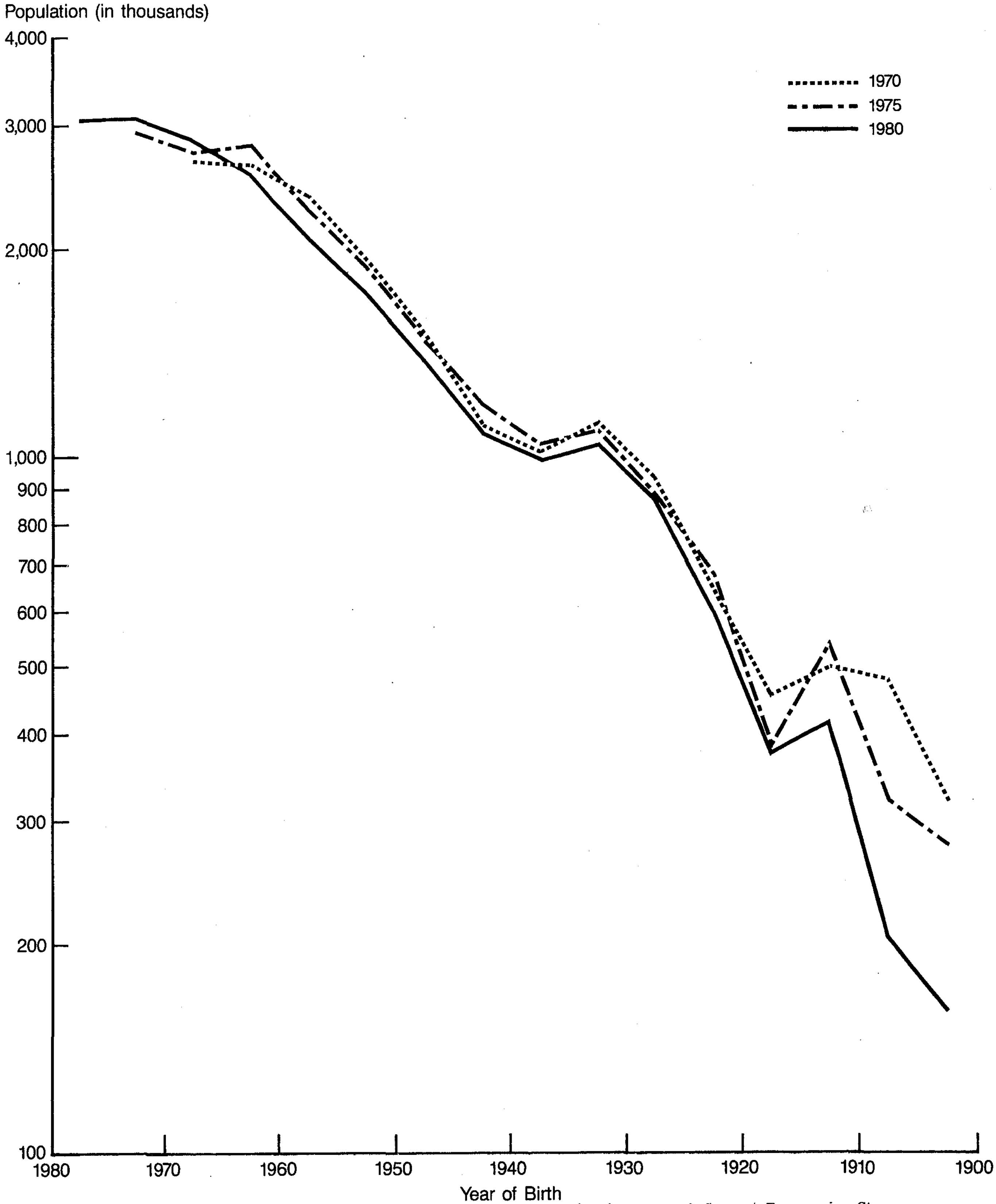
Source: Yemen Arab Republic (1980), table 1.

Figure 5-10. CENSUS POPULATION OF ARGENTINA, MALES BY SINGLE YEARS OF AGE, 1980



Source: Argentina (1982), table G.2.

Figure 5-11. GRAPHICAL COHORT ANALYSIS OF CENSUS ENUMERATIONS OF MALES IN THE TURKISH CENSUSES OF 1970, 1975, AND 1980



Source: State Institute of Statistics, Census of Population, Social and Economic Characteristics of the Population, table 8 (for 1970 and 1975) and table 14 (for 1980).

historical events rather than census errors (although some degree of census error is suggested by the fact that the lines for the three censuses cross during those periods). The implication of this is that these distortions should not be "smoothed" when using the census age distributions for population projections or other demographic analyses since they reflect actual characteristics of the population.

2.2 *Age and sex ratios*

The assessment of the "reasonableness" of census age-sex distributions may be approached in a more quantitative fashion through the examination of age and sex ratios in the enumerated census population. The diagnostic value of these ratios lies in their predictable patterns in human populations. Accordingly, significant departures from ratio values expected in the absence of measurement error which cannot be explained by extraordinary or distorting factors indicate census error.

Age ratios provide a measure of the "smoothness" of the age distribution of the population over a restricted age range. In the absence of sharp swings in fertility or mortality, significant levels of migration, or other distorting factors, the enumerated size of a particular cohort should be approximately equal to the average size of the immediately preceding and subsequent cohorts. In other words, the ratio of the census count for a particular cohort to the average of the counts for the adjacent cohorts should be approximately equal to 1.0 (or 100 if multiplied by a constant of 100). Significant departures from this "expected" ratio indicate either the presence of error in the census enumeration or of other factors noted above.

Sex ratios are defined as the ratio of the enumerated male population to the enumerated female population in a particular age group (usually multiplied by a constant to produce a measure based upon 100 - that is, the number of males per 100 females). The sex ratio at birth for most populations lies between 102 to 107, but due to higher mortality rates for males than females in most populations, the ratio declines gradually among successively older segments of the population. Accordingly, significant variations from this pattern are suggestive either of census errors or of the presence of distorting factors such as sex-selective migration, periods of high or changing sex-selective mortality, or in some populations higher female than male mortality.

The uses of age and sex ratios for census evaluation purposes are illustrated here using the data from the censuses of Yemen and Argentina (see figures 5-5 and 5-6). Figure 5-12 shows the age ratios for 5-year age categories by sex calculated from these two censuses. The age ratios were calculated using the formula:

$$(5.1) \quad {}_5AR_x = \frac{{}_5P_x}{\frac{1}{3}\left({}_5P_{x-5} + {}_5P_x + {}_5P_{x+5}\right)} \times 100$$

Where:

${}_5AR_x$ = The age ratio for the age category x to $x+4$

${}_5P_x$ = The enumerated population in the age category x to $x+4$

${}_5P_{x-5}$ = The enumerated population in the adjacent lower age category (ages $x-5$ to $x-1$)

${}_5P_{x+5}$ = The enumerated population in the adjacent higher age category (ages $x+5$ to $x+9$)

The irregularities in the age distributions noted in the graphical analyses of the two censuses are clearly reflected by the age ratios shown in figure 5-12. The ratios for the Yemen census reveal a larger number of significant departures from the expected ratio value of 100 than the Argentine census. The general pattern in the Yemen census is

Figure 5-12. AGE RATIOS FOR FIVE-YEAR AGE CATEGORIES BY SEX: 1975 CENSUS OF YEMEN AND 1980 CENSUS OF ARGENTINA

Age	YEMEN		ARGENTINA	
	Males	Females	Males	Females
0 to 4 years......	-	-	-	-
5 to 9 years.......	117.2	115.9	98.5	98.5
10 to 14 years.....	97.3	87.9	97.3	97.0
15 to 19 years.....	85.9	94.7	100.2	99.9
20 to 24 years.....	81.1	90.1	99.3	100.2
25 to 29 years.....	107.6	109.1	100.7	100.9
30 to 34 years.....	97.0	115.2	101.9	101.5
35 to 39 years.....	105.1	99.2	98.8	98.8
40 to 44 years.....	105.2	107.2	97.6	97.2
45 to 49 years.....	86.6	82.3	100.6	98.8
50 to 54 years.....	124.6	129.0	102.4	104.2
55 to 59 years.....	65.8	60.1	103.5	101.9
60 to 64 years.....	144.6	153.9	94.7	95.7
65 to 69 years.....	64.8	56.9	104.0	104.8
70 to 74 years.....	135.6	148.9	97.6	99.0
75 to 79 years.....	-	-	-	-

Source: Derived from figures 5-5 and 5-6.

for the magnitude of the departures from a ratio of 100 to grow larger with the increasing age of the population. This is especially true after age 40, due to a substantial preference among the population aged 40 and above for ages ending in the digit 0 (and to a lesser extent the digit 5) observed in figures 5-7 and 5-9. The fluctuations in the age ratios for the population aged 60 and above in the Argentine census are also suggestive of somewhat greater census error for the older than for the younger population.

Two other points worthy of note in connection with the age ratios from Yemen are the high ratios for the 5 to 9 cohort and the low ratios for the 15 to 19 and 20 to 24 cohorts. The very large census count for the 5 to 9 cohort, especially in relation to the 0 to 4 cohort, again suggests the presence of significant levels of underenumeration of young children and/or age misreporting resulting in a net transfer of population into the 5 to 9 category in the Yemen census. In actual application, further investigation of the consistency of the enumerated number of children aged 0 to 4 with the prevailing level of fertility and the number of women of childbearing age (ages 15 to 49) would have to be undertaken in order to assess the relative magnitude of coverage and content errors for these cohorts.

The low ratios for the 15 to 19 and 20 to 24 cohorts suggest the presence of some combination of census error (see figures 5-7 and 5-9 for evidence of age heaping) and labor migration to other middle eastern countries (particularly Saudi Arabia) which has characterized Yemen in recent years. Higher mortality associated with the civil war (1968-1975) may also have contributed to the lower age ratios for this segment of the population. However, the fact that the male age ratios in the 15 to 19 and 20 to 24 age categories are significantly lower than those for females supports the interpretation that labor migration is an important contributing factor in the unexpectedly small cohorts in this age range (at least for males).

By way of contrast, the age ratios calculated from the 1980 Argentine census (figure 5-12) show relatively minor deviations from the expected ratio value of 100 up to approximately age 60, at which point the ratios become somewhat more erratic. The pattern of age ratios in the Argentine census provides no evidence of significant systematic error in the census enumeration (of the population aged 60 and under, at least), although the relatively low ratios for the 5 to 9 and 10 to 14 cohorts suggest underenumeration and/or age misreporting.

Figure 5-13. CENSUS OF POPULATION BY AGE AND SEX AND SEX RATIOS: 1975 CENSUS OF YEMEN AND 1980 CENSUS OF ARGENTINA

Age	YEMEN			ARGENTINA		
	Males	Females	Ratio	Males	Females	Ratio
Total.......	2,152,554	2,364,637	91.0	13,755,983	14,191,463	96.9
0 to 4 years.....	392,947	393,969	99.7	1,640,005	1,600,836	102.4
5 to 9 years.....	432,284	401,496	107.7	1,407,213	1,376,543	102.2
10 to 14 years...	281,036	243,576	115.4	1,240,209	1,215,962	102.0
15 to 19 years...	153,427	185,793	82.6	1,173,841	1,167,647	100.5
20 to 24 years...	101,488	159,447	63.6	1,099,810	1,124,347	97.8
25 to 29 years...	120,335	185,731	64.8	1,050,065	1,074,218	97.8
30 to 34 years...	113,706	165,673	68.6	979,816	994,377	98.5
35 to 39 years...	117,589	143,951	81.7	855,293	869,013	98.4
40 to 44 years...	104,442	125,601	83.2	772,913	775,702	99.6
45 to 49 years...	75,741	81,976	92.4	748,046	748,741	99.9
50 to 54 years...	82,106	91,371	89.9	709,825	749,808	94.7
55 to 59 years...	39,915	39,211	101.8	620,972	659,561	94.1
60 to 64 years...	59,834	65,237	91.7	469,993	532,126	88.3
65 to 69 years...	24,396	22,746	107.3	397,567	476,239	83.5
70 to 74 years...	28,641	32,041	89.4	279,279	355,241	78.6
75 or older......	24,647	26,818	92.0	311,136	471,102	66.0

Source: Derived from figures 5-5 and 5-6.

The sex ratios from the Yemen and Argentine censuses shown in figure 5-13 provide additional insight into likely errors in these data. The ratios were calculated using the following formula:

$$(5.2) \quad {}_5SR_x = \frac{\text{number of males enumerated in the age category } x \text{ to } x + 4 \text{ years}}{\text{number of females enumerated in the age category } x \text{ to } x + 4 \text{ years}} \times 100$$

While the expected pattern of smoothly declining sex ratios with increasing age in the population is for the most part observed in the Argentine census, the sex ratios for Yemen are quite volatile. The low ratio for the 0 to 4 age group and the high ratio for the 5 to 9 age group suggests the possibility of substantial sex-selective age misreporting such that a significant number of male children aged 0 to 4 years were reported as being ages 5 to 9 years in the census. The high ratio for the 10 to 14 cohort is also suggestive of the presence of sex-selective coverage and/or age misreporting errors. On the other hand, the very low ratios for the three cohorts beginning with the 20 to 24 cohort are supportive of the interpretation that significant male labor migration has affected the age distribution of the population of Yemen. Finally, the fact that the ratios for older cohorts (ages 45 and above) are quite volatile and never fall much below 90 are indicative of significant levels of coverage and content (age reporting) error in the census of Yemen.

The only notable distortion in the progression of sex ratios in the Argentine census is the relatively low ratios for the 20 to 24 and 25 to 29 cohorts and the pattern of slightly increasing sex ratios over the 20 to 49 age range. This likely reflects some combination of sex-selective migration and possibly differential census coverage by sex (that is, females being enumerated more completely than males). Overall, however, the degree of distortion is relatively small.

The effects of age- and sex-selective international migration on the age-sex distribution of Yemen illustrate a frequently encountered problem when using age- and sex-ratio analysis to evaluate census results. When ratios are to be used to evaluate census results for sub-areas of a country, this problem is further compounded by age- and sex-selective internal migration. This is illustrated by the sex ratios for urban and rural areas calculated from the 1980 census of Argentina (figure 5-14). Beginning with the 15 to 19 cohort, the sex ratios in urban areas are consistently below 100, while those in rural areas are consistently above 100. These ratios reflect a pattern of female-selective migration from rural to urban areas which generally characterizes Latin American countries and also has been observed in several Asian countries (Shaw, 1974). The implication of this is that in instances in which sex and/or age ratios tend to be distorted by migration, alternative "expected" ratios must be derived for the purposes of census evaluation.

Figure 5-14. SEX RATIOS OF THE CENSUS POPULATIONS IN URBAN AND RURAL AREAS IN ARGENTINA, 1980

Age	Total	Urban	Rural
All ages......	96.9	93.6	114.9
0 to 4 years......	102.4	102.4	102.6
5 to 9 years......	102.2	101.9	103.3
10 to 14 years....	102.0	100.1	109.6
15 to 19 years....	100.5	96.3	122.2
20 to 24 years....	97.8	95.1	113.5
25 to 29 years....	97.8	94.5	117.5
30 to 34 years....	98.5	95.1	119.9
35 to 39 years....	98.4	95.0	119.4
40 to 44 years....	99.6	95.6	123.9
45 to 49 years....	99.9	95.5	128.3
50 to 54 years....	94.7	90.1	125.9
55 to 59 years....	94.1	89.3	129.7
60 to 64 years....	88.3	82.8	127.8
65 to 69 years....	83.5	77.9	126.9
70 to 74 years....	78.6	73.2	124.1
75 years or older.	66.0	61.9	100.6

Source: Argentina (1980), table G.2.

One possibility for deriving "expected" sex-ratios in a current census is to use the sex-ratio of the population enumerated in a previous census as the basis for evaluation. The most general use of sex-ratios in successive censuses would entail the assessment of the plausibility of the ratios of the total populations enumerated. Unless affected by errors or significant shifts in the sex-selectivity of migration or mortality, the sex-ratio of the enumerated total population should be relatively stable from census to census.

As an illustration, figure 5-15 shows the sex-ratios of enumerated populations from the most recent censuses of five countries which have conducted three or more censuses. On a comparative basis, the sex-ratios in the Indian and Indonesian censuses would appear to exhibit the greatest degree of stability from census to census. To some extent, this may reflect the lower volume of international migration to and from these countries in comparison with the other countries considered. The Egyptian and Philippine censuses, on the other hand, show somewhat greater variability from census to census. In the Egyptian case, the relatively high sex ratio of the population enumerated in the 1976 census in particular stands out from the sex-ratios in the four previous censuses (which vary within a narrower range) and would warrant further investigation in actual application. Some of the changes in the sex ratios in successive censuses in the Philippines are likely attributable to the changing sex-selectivity of emigration from the Philippines from female-dominated migration to the United States and

Figure 5-15. SEX RATIOS OF THE POPULATION ENUMERATED IN THE MOST RECENT CENSUSES OF SELECTED COUNTRIES

Country	Year of Census	Sex Ratio
Egypt........	1927	99.1
	1937	100.2
	1947	98.1
	1960	101.2
	1976	104.1
India........	1941	105.8
	1951	105.7
	1961	106.3
	1971	107.5
	1981	106.9
Indonesia....	1961	97.3
	1971	97.2
	1980	98.8
Pakistan.....	1951	116.8
	1961	115.8
	1972	114.9
	1981	110.4
Philippines..	1960	101.8
	1970	99.0
	1975	102.3
	1980	100.7

Source: Egypt (1979); India (1975, 1983); Indonesia (1963, 1975, and 1983); Pakistan (U.S. Bureau of the Census, 1984: table 4); Philippines (1963, 1974, 1978, and 1983).

Canada in the 1960's and early 1970's to a significant volume of male labor migration to middle eastern countries in the mid to late 1970's. Male labor migration is also likely to have been a significant contributing factor in the decline in sex-ratios in successive censuses in Pakistan, which otherwise would appear to have been relatively consistent over time.

Comparisons of sex-ratios from successive censuses can also be helpful in detecting problems in particular age ranges or for particular cohorts. With respect to age categories, the expected relationship of sex-ratios in any given age range in successive censuses would be one of relative stability in the absence of census errors, changes in the sex-selectivity of migration, or other distorting factors. Significant fluctuations from this pattern which cannot be plausibly explained by one or more distorting factors are likely to indicate variations in coverage or accuracy of age reporting on a sex-selective basis from census to census.

Similarly, in the absence of census errors or distorting factors, changes in the sex-ratios of birth cohorts from census to census should be consistent with sex differentials of mortality under the prevailing mortality regime(s). In countries where male mortality rates exceed those of females, the expected pattern would be for the sex-ratio for each cohort to decline from census to census as the cohort ages. The opposite pattern would be expected in countries where female mortality rates exceed those for males.

The use of cohort-specific sex-ratios from successive censuses is illustrated in figure 5-16, which shows sex-ratios for 5-year age groups from the four most recent censuses of Pakistan. Looking first at the degree of consistency of sex-ratios across censuses for fixed age categories, several observations may be made. First, the consistency of sex-ratios appears to be somewhat greater in the three most recent censuses than when the 1951 census is considered, suggesting some degree of differential coverage/accuracy in the 1951 census in comparison with the 1961, 1972, and 1981 censuses (shown graphically in figure 5-17). Second, even within the more homogeneous group of censuses, significant fluctuations in sex-ratios for particular age groups are observed (for example, the 10 to 14 and 50 to 54 age groups). Third, the decline in the sex-ratios in the age groups in the 25 to 49 range, particularly between the 1972 and 1981 censuses, is consistent with expectations in view of the signficant volume of male labor

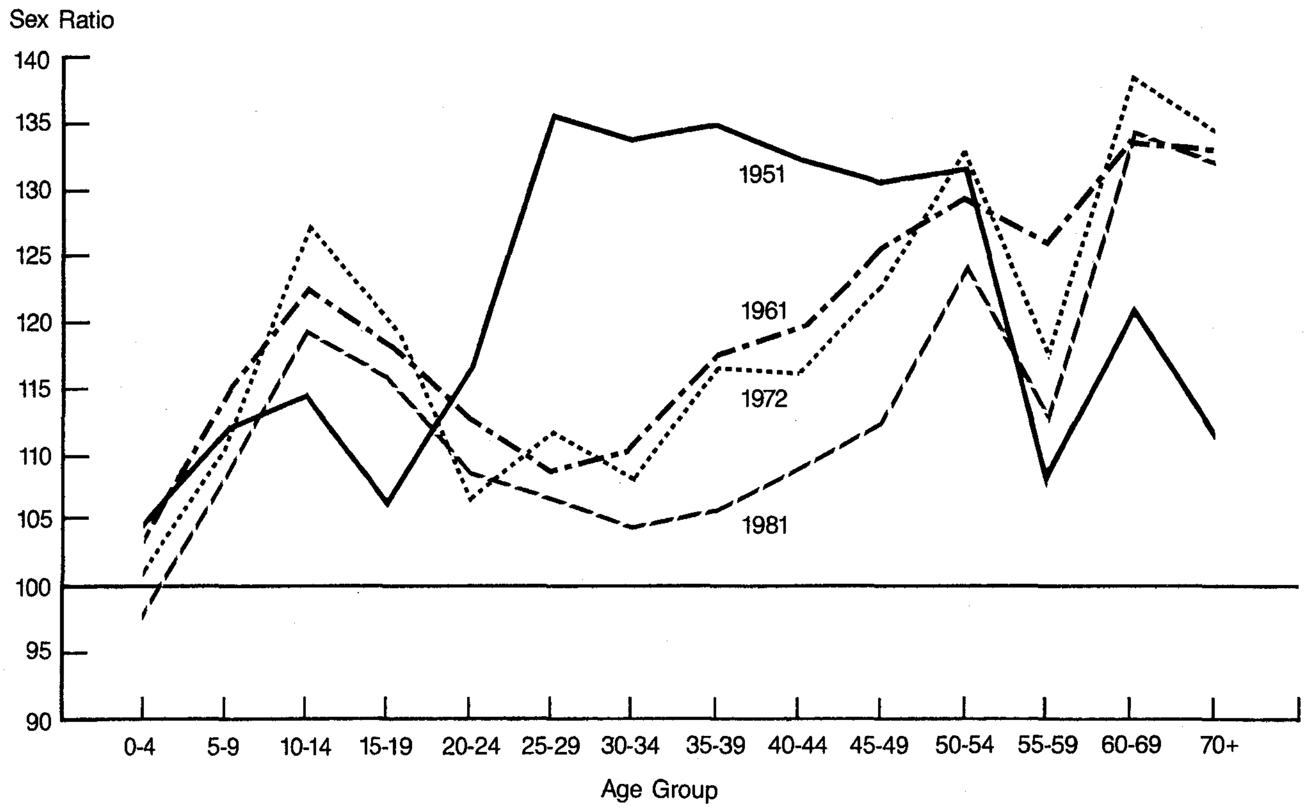
Figure 5-16. SEX RATIOS BY AGE GROUP FOR THE POPULATION ENUMERATED IN THE 1951, 1961, 1972, AND 1981 CENSUSES OF PAKISTAN

Age	1951	1961	1972	1981
All ages......	116.8	115.8	114.9	110.4
0 to 4 years.......	104.6	103.0	100.8	97.7
5 to 9 years.......	112.0	114.6	110.4	108.4
10 to 14 years.....	114.4	122.6	127.0	119.3
15 to 19 years.....	106.1	118.5	120.1	115.9
20 to 24 years.....	116.5	112.9	106.3	108.6
25 to 29 years.....	135.1	108.9	111.6	106.6
30 to 34 years.....	133.9	110.7	108.1	104.4
35 to 39 years.....	134.9	117.5	116.4	105.8
40 to 44 years.....	132.2	119.4	116.1	109.2
45 to 49 years.....	130.2	125.5	122.9	112.4
50 to 54 years.....	131.4	129.4	132.7	124.0
55 to 59 years.....	108.6	125.8	118.2	113.3
60 to 69 years.....	121.2	133.7	138.5	134.2
70 years or older..	111.7	132.9	134.6	131.8

Source: U.S. Bureau of the Census (1984), table 4.

migration from Pakistan to the middle east. Finally, the steady decline in the sex-ratio of the population aged 0 to 4 years in each successive census is somewhat puzzling in view of the fact that the sex-ratios of the population in the 5 to 9 age group also declined in successive censuses after 1961. If increasing levels of exaggeration of ages of male children were responsible for the declining sex-ratio of the population aged 0 to 4 years, it would be anticipated that the sex-ratio of the 5 to 9 age group would have increased. The fact that they did not suggests either that the sex ratio at birth has declined in Pakistan or that female children were enumerated increasingly more completely in each successive census relative to male children.

Figure 5-17. SEX RATIOS BY AGE GROUP FOR THE POPULATION ENUMERATED IN THE 1951, 1961, 1972, AND 1981 CENSUSES OF PAKISTAN



Source: Derived from figure 5-16.

Significant variations in the sex-ratios of birth cohorts from census to census are also apparent in the data displayed in figure 5-18. Given the higher levels of female than male mortality in Pakistan, the expected pattern would be for the sex-ratio of the population enumerated of each cohort to increase in successive censuses. This pattern is observed for relatively few cohorts, however, suggesting the presence of census error and/or distorting factors. Labor migration might partially account for the declining sex ratios of the cohorts under age 15 at the time of the 1951 census.

2.3 *Summary indices of error in age-sex data*

A number of measures have been developed to summarize the magnitude of observed discrepancies in census age-sex distributions in the form of a single index value. Among the more widely-used are the United Nations Age-Sex Accuracy Index (United Nations, 1955), Whipple's Index (1921), Myers' Blended Index (1940), and an index recommended by the United Nations which provides a measure of the relative importance of age overstatement and understatement in accounting for age heaping (United Nations, 1955).

Figure 5-18. SEX RATIOS OF BIRTH COHORTS AS ENUMERATED IN SUCCESSIVE CENSUSES IN PAKISTAN, 1951 TO 1981

Age of birth cohort in the 1951 Census	1951	1961	1972	1981
0 to 4 years.......	104.6	122.6	106.3	104.4
5 to 9 years.......	112.0	118.5	111.6	105.8
10 to 14 years.....	114.4	112.9	108.1	109.2
15 to 19 years.....	106.1	108.9	116.4	112.4
20 to 24 years.....	116.5	110.7	116.1	124.0
25 to 29 years.....	135.1	117.5	122.9	113.0
30 to 34 years.....	133.9	119.4	132.7	134.2
35 to 39 years.....	134.9	125.5	118.2	131.8
40 to 44 years.....	132.2	129.4	138.5	-
45 to 49 years.....	130.2	125.8	134.6	-
50 to 54 years.....	131.4	133.7	-	-
55 to 59 years.....	108.6	132.9	-	-

Source: Derived from figure 5-16.

The United Nations Age-Sex Accuracy Index is essentially an extension of age- and sex-ratios analysis in which the deviations of the observed from the expected age and sex ratios for each 5-year age group are combined into a single score. A low index value is indicative of relatively minor deviations from the expected age and sex ratios, while high index values indicate more serious distortions in the age-sex distribution.

The Whipple's and Myers' indices measure the magnitude of a particular type of error in census age data - preference for ages ending in particular digits (age heaping). Whipple's Index measures the extent of heaping on ages ending in the digits 0 to 5 on a scale with a range from 0 to 500, where 0 represents total "avoidance" of these digits, 100 represents a uniform distribution across ages ending in each of the digits 0 to 9 (that is, no age heaping), and 500 represents the extreme case of digit preference in which the entire census population is recorded at ages ending in 0 or 5.

Myers' Blended Index is conceptually similar to Whipple's Index, except that the index considers preference (or conversely, avoidance) of ages ending in each of the digits 0 to 9 in deriving an overall age-accuracy score. The theoretical range of the Myers' Index is from 0 to 90, where an index value of 0 indicates no age heaping and 90 indicates the extreme case where all recorded ages end in the same digit.

The United Nations has recommended a procedure to examine the causes of age heaping in a more detailed manner. Specifically, the measure assesses whether age overstatement or age understatement are more important causes of observed heaping on particular digits. For each terminal digit 0 or 5 (denoted age x) where digit preference is observed, the procedure entails the calculation of the ratio of the population enumerated in the next younger age $(x-1)$ to the population count for the next older age $(x+1)$. A ratio value of 1.0 would indicate that age understatement and overstatement contributed equally to the observed level of age heaping on age x. A ratio value of greater than 1.0 would indicate that age understatement was a more important determinant of age heaping on age x than age overstatement, while a ratio of less that 1.0 would have the opposite interpretation.

The application of these procedures is not illustrated in this manual for two primary reasons. First, the procedures are well-documented in other widely-available reference sources (U.S. Bureau of the Census, 1975; United Nations, 1955). Second, these procedures are summary measures of error in census age and sex data and, as such, are not an adequate substitute for detailed inspection of the data as described above. While these procedures are useful as summary measures or for comparative purposes, they generally do not provide any insight into patterns of error in the data that cannot be obtained through graphical and ratio analyses of the data. Where a summary measure is desired after detailed inspection of the data, however, the reader is referred to the sources cited above for detailed computational procedures.

2.4 *Uses and Limitations*

The assessment of the reasonableness of the age and sex distribution of the population enumerated in a census is typically the first step taken in evaluating a census by means of demographic methods. Such analyses serve several useful purposes. First, they provide a quick and inexpensive indication of the general quality of the data. Second, they provide evidence on the specific segments of the population in which the presence of error is likely. Finally since age-sex distributions reflect levels and trends in fertility, mortality, migration, and extraordinary events experienced in the country's past, such analyses can provide "historical" information which may be useful for interpreting the results of evaluation studies based upon other methods and in determining how the census data should be adjusted for use in subsequent demographic analyses.

The major limitation of age-sex distribution analysis is that, generally speaking, it is not possible to derive separate numerical estimates of the magnitude of coverage and content error on the basis of such analyses alone. While it is often possible, as illustrated above, to discern that particular types of errors are likely to have affected the census counts for particular segments of the population with a fairly high degree of confidence, estimates of coverage error from other sources (PES, for example) often are required to verify these observations and permit the development of separate numerical estimates of the degree of census coverage and content error.

3. *STABLE POPULATION ANALYSIS OF AGE DISTRIBUTIONS*

Another approach for assessing the plausibility of a census age and sex distribution is based upon the comparison of the recorded age distribution (for each sex separately) with the age distribution of an appropriately chosen stable population. As indicated in chapter 4, the age distribution of a population which has been subject to constant levels of fertility and mortality and no international migration will, given a sufficiently long period of time, evolve into a constant distribution which is independent of the initial age distribution.

3.1 *Basis of method*

The constant age distribution of a stable population, known as a stable age distribution, is determined fully by the prevailing constant levels of fertility and mortality (assuming the absence of migration) and has the following form:

(5.3) $c(x) = b\ \ell(x)\ exp\ (-rx)$

Where:

$c(x)$ = the infinitesimal proportion of the stable population at exact age x

b = the constant birth rate

r = the constant rate of natural increase

$\ell(x)$ = the probability of survival from birth to age x

From equation (5.3), the proportion of the population under age y, $C(y)$, and the birth rate, b, of a stable population may be derived as shown in equations (5.4) and (5.5), respectively.

(5.4) $C(y) = \int_0^x b\ \ell(x)\ exp\ (-rx)\ dx$

(5.5) $b = \left[\int_0^w \ell(x)\ exp\ (-rx)\ dx\right]^{-1}$

Where:

w = The highest age attainable in the population, and all other terms are as defined above

Equations (5.4) and (5.5) permit the calculation of an "expected" stable age distribution for an actual population when two of three parameters b, r, or $\ell(x)$ are known or can be estimated. In populations which have experienced approximately constant levels of fertility, negligible levels of net international migration, and mortality levels which have been either constant or have declined only recently, a stable age distribution derived on this basis provides a meaningful standard against which to compare a recorded census age distribution.

In actual application, few (if any) populations are genuinely stable. Even in the absence of wide swings in fertility and mortality levels, periodic fluctuations can produce age distributions which do not conform precisely to those anticipated through the application of stable population theory. This problem is compounded by the fact that the parameters b, r, and/or $\ell(x)$ used to derive a stable age distribution for census evaluative purposes are often themselves indirect estimates which are, as such, subject to error.

Despite these problems, previous experience has shown stable population anaysis to be a useful tool for evaluating age distributions in populations which have been subject to approximate stability or low-level fluctuations in vital rates and migration (United Nations 1983). However, an underlying consideration in meaningful application of stable population theory to demographic estimation or census evaluation is that, because of the problems cited above, estimates finally accepted should be based upon central tendencies in a series of estimates identified using different pairs of values of the parameters utilized in the analyses. This point is illustrated in the examples provided below and discussed in greater detail in United Nations (1967 and 1983).

The use of stable population methods is based upon the various model stable populations described in chapter 4. This summary description and the references cited therein should be consulted in connection with the illustrative examples provided below.

The following data are required for stable population analysis:

(1) The census count of population (which is to be evaluated by single years of age or 5-year age groups, by sex; and

(2) Estimates of two of the following parameters: (a) the growth rate r in the population; (b) the birth rate b; and (c) the probability of surviving from birth to age x; that is, the $\ell(x)$ function of a life-table. An estimate of expectation of life at birth e, may be used to select a model life-table to represent mortality conditions in the population under study.

Procedures for deriving indirect estimates of these parameters are summarized in section 4.3.

3.2 *Computational procedure*

For the procedure involving direct comparison between the recorded and stable age distribution, the computational steps described below are required.

3.21 Step 1: Calculation of the proportional age distribution of the census population.--The initial step in the procedure is to calculate the proportional distribution $c(x)$ of the enumerated census population (for each sex) by 5-year age categories. If ${}_5N_x$ represents the number of persons aged x to $x + 4$ enumerated in the census and N is the total population enumerated, the proportion in each category x is calculated as:

(5.6) $$c(x) = {}_5N_x / N \times 100$$

3.22 Step 2: Selection of a model stable age distribution.--The model stable age distribution against which the recorded census distribution is to be compared is identified by calculating the proportions in each age group in the model stable population, which shall be denoted $c(x)'$, which corresponds to the known or estimated values of two of the parameters r, b, or in $\ell(x)$ the population under study. If, for example, r and $\ell(2)$ were to be used to select the model stable population, values of $c(x)'$, that is those for the stable age distribution, would be obtained by interpolating between the printed values of $\ell(2)$ and r in the model life tables and stable populations to identify the implied stable population, and then using the same interpolation factors to calculate the corresponding values of $c(x)'$ in that stable population. This two-way interpolation procedure is used irrespective of which set of parameters is used to identify a model stable population. The computations involved are illustrated in the examples below.

3.23 Step 3: Comparison of the recorded and stable age distribution.--The final step in the procedure is to compare the age distribution for each sex recorded in the census with the stable age distribution implied by the two demographic parameters for the population. A useful measure of the degree of departure of the recorded age distribution from that expected on the basis of the stable population theory is the ratio of the census to the stable populations proportions in each age category, or $c(x)/c(x)'$. A ratio of 1.0 would indicate exact correspondence between the observed and expected age distributions. Ratios of less than 1.0 in a particular age category would indicate a lower than predicted census count in that category and would be suggestive of net underenumeration and/or a net transfer of persons out of the category through age misreporting (content error). A ratio of greater than 1.0 would have the opposite interpretation.

3.3 *Examples of stable population theory applications*

Two examples are presented in this section in order to illustrate somewhat different applications of stable population theory for census evaluation purposes. The first application entails the direct comparison of a recorded census age distribution and a stable age distribution selected on the basis of two of the parameters from equation (5.4).

The second procedure described does not directly compare the recorded and stable age distributions, but rather assesses the consistency of the stable implied birth rate at different points in the cumulative age distribution, *b(x)*, derived on the basis of the recorded census age distribution and an arbitrarily chosen level of mortality. In the application, inconsistencies in the implied birth rate at different points in the age distribution from a constant value are suggestive of the errors in the recorded age distribution (assuming, of course, that the population was in fact, approximately stable).

3.31 1970 Census of Thailand.--As an illustration, the procedure is first applied to the 1970 Census of Thailand to assess the plausibility of the distribution of the enumerated population by age and sex (shown in figure 5-19).

As described above, the first step in the computational procedure is to transform the census counts of persons by age and sex into a proportional age distribution (for each sex). The figures shown in figure 5-19 were used to derive the proportional age distributions for males and females (that is, the proportions in each age category) shown in columns (1) and (4) of figure 5-20.

The next step in the computational procedure is to select a model stable age distribution consistent with demographic conditions in Thailand during the 1960-1970 period for use as a standard in evaluating the recorded age distribution in the 1970 census. In the case of Thailand, sufficient information is available for the 1960-1970 period such that any combination of two of the parameters r, b, or $\ell(x)$ may be used in selecting a model stable population. Following Arnold and Phananiramai (1975), the intercensal growth rate, r, and the estimated level of mortality during the intercensal period were used in this illustration.

The growth rate, r, was derived using the census counts from the 1960 and 1970 censuses and solving the following equation for r.

Figure 5-19. POPULATION ENUMERATED IN THE 1970 CENSUS OF THAILAND, BY AGE AND SEX

Age	Total	Male	Female
All ages	34,397.3	17,124.0	17,273.3
0 to 4 years	5,666.4	2,866.6	2,799.8
5 to 9 years	5,291.6	2,682.6	2,609.0
10 to 14 years	4,568.0	2,312.5	2,255.5
15 to 19 years	3,722.3	1,834.5	1,887.8
20 to 24 years	2,686.7	1,323.3	1,363.4
25 to 29 years	2,244.3	1,099.5	1,144.8
30 to 34 years	2,127.0	1,048.6	1,078.4
35 to 39 years	1,913.0	954.2	958.8
40 to 44 years	1,542.6	775.3	767.3
45 to 49 years	1,198.1	599.9	598.2
50 to 54 years	963.2	472.8	490.4
55 to 59 years	791.0	388.8	402.2
60 to 64 years	625.8	301.2	324.6
65 to 69 years	452.4	213.2	239.2
70 years or older	604.9	251.0	353.9

Source: Thailand (1973).

(5.7) $P_1 = P_0 \; exp \; (rt)$

Inserting the respective census population counts and the length of the intercensal interval (t = 10 years) into the equation, we have:

(5.8) $34{,}397{,}300 = 26{,}432{,}000 \; exp \; (r \cdot 10)$

Taking **natural** logarithms of both sides of equation (5.8) and rearranging terms yields:

$$r = \frac{\ell n \; (34{,}397{,}300 \div 26{,}432{,}000)}{10}$$

$$\cong .03$$

Figure 5-20. COMPARISON OF AGE DISTRIBUTION OF THE POPULATION ENUMERATED IN THE 1970 CENSUS OF THAILAND AND A STABLE AGE DISTRIBUTION, BY SEX

Age	Males			Females		
	Enumerated Population (1)	Stable Population[1] (2)	Ratio (3)=(1)÷(2)	Enumerated Population (4)	Stable Population[1] (5)	Ratio (6)=(4)÷(5)
All ages...	100.0	100.0	-	100.0	100.0	-
0 to 4 years....	16.7	17.3	0.97	16.2	17.0	0.95
5 to 9 years....	15.7	14.4	1.09	15.1	14.2	1.06
10 to 14 years..	13.5	12.3	1.10	13.1	12.1	1.08
15 to 19 years..	10.7	10.5	1.02	10.9	10.3	1.06
20 to 24 years..	7.7	8.8	0.88	7.9	8.8	0.90
25 to 29 years..	6.4	7.5	0.85	6.6	7.4	0.89
30 to 34 years..	6.1	6.3	0.97	6.2	6.2	1.00
35 to 39 years..	5.6	5.3	1.06	5.6	5.3	1.06
40 to 44 years..	4.5	4.4	1.02	4.4	4.4	1.00
45 to 49 years..	3.5	3.6	0.97	3.5	3.7	0.95
50 to 54 years..	2.8	2.9	0.97	2.8	3.0	0.93
55 to 59 years..	2.3	2.3	1.00	2.3	2.4	0.96
60 to 64 years..	1.8	1.7	1.06	1.9	1.9	1.00
65 to 69 years..	1.3	1.2	1.08	1.4	1.4	1.00
70 years or older.....	1.5	1.5	1.00	2.0	1.9	1.05

Sources: Thailand (1973): Coale and Demeny (1966)
[1]*Stable age distribution with "West" mortality, level 17, and r=.03 (see text for further details on method of selection of this particular stable age distribution).*

Information on the level of mortality during the intercensal period was provided by the Survey of Population Change (Thailand, 1969) conducted during the 1964-1967 period and estimates derived by the U.S. Bureau of the Census (1978). These data suggest an average expectation of life at birth, e_0 for females during the intercensal period of approximately 60 years. Based upon this estimate, level 17 of the Coale and Demeny (1966) "West" model life tables was deemed to be representive of mortality conditions in Thailand during the 1960-1970 intercensal period (female e_0 = 60 years; male e_0 = 56.5 years).

The age distribution of the Coale and Demeny stable population corresponding to level 17 "West" mortality with an annual growth rate, r, of three percent (.03) was chosen accordingly as the "standard" against which to evaluate the recorded age distribution in the 1970 census. These stable age distributions for males and females are shown in columns (2) and (5) respectively, of figure 5-20. In this particular example, interpolation is not required since the estimated r and e_0 parameters for Thailand are printed in the Coale-Demeny model life tables and stable populations. In other cases, as in the second example presented below, it will be necessary to interpolate between the printed values in the model life table and stable populations to derive appropriate values for the population under study.

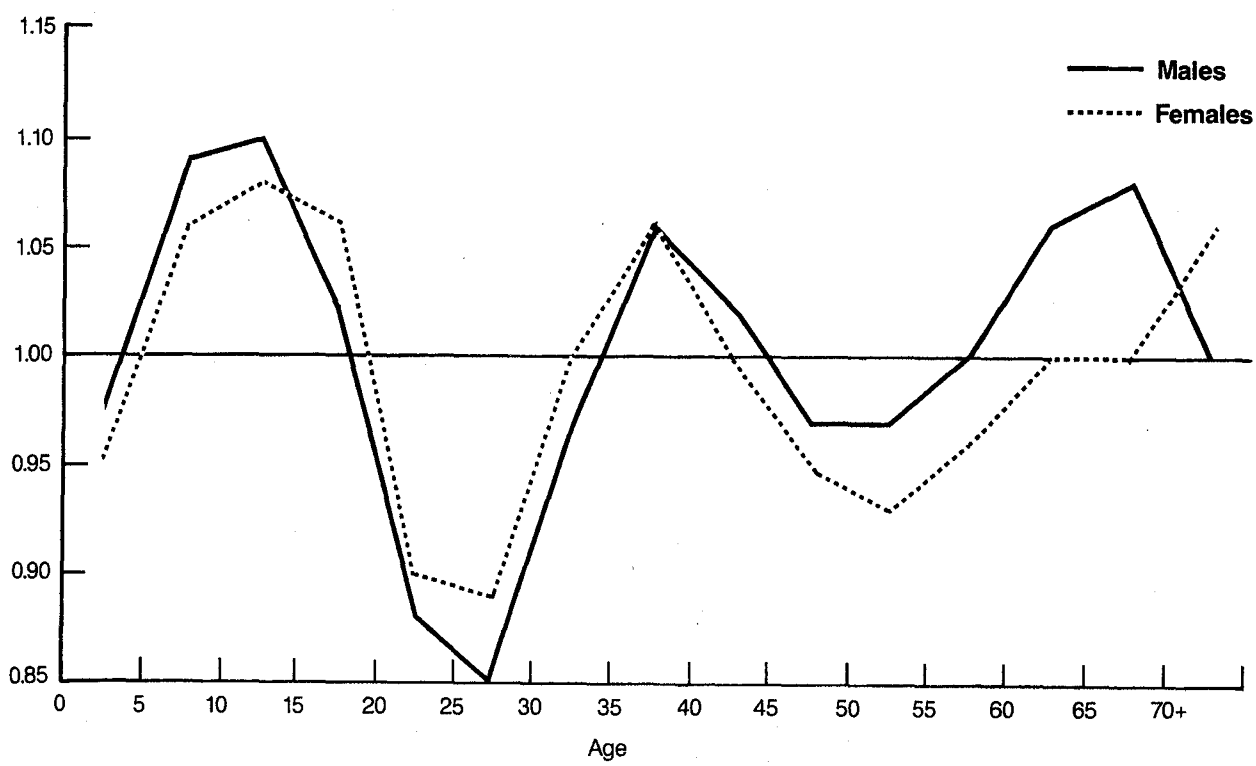
The degree of correspondence between the recorded and stable age distributions in the case of the 1970 census of Thailand is indicated by the ratios of the recorded to the "expected" proportions in each age category shown in columns (3) and (6) of figure 5-20 and displayed graphically in figure 5-21.

These results suggest the following irregularities: (a) an underenumeration of children aged 0 to 4 years; (b) a relatively high proportion of the census population under age 15; (c) a relatively low proportion of the population in the 20 to 34 age range; (d) somewhat more modest departures from the stable age distribution at ages 35 to 59; and (e) somewhat erratic counts above age 60 for males and 65 for females.

If the population of Thailand had been genuinely stable, these results would be indicative of significant levels of coverage and/or content error in the 1970 census. In fact, the major distortions in the 1970 census age distributions shown in figure 5-21 are explainable to some extent by changes in fertility levels since World War II. The deficit of population ages 25 to 34, for example, is likely the result of lower fertility levels during the war years (see Bourgeois-Pichat, 1974), although some census underenumeration (particularly of males) in this age range is also likely. Similarly, the relatively high proportion of the population under age 20 may be attributable to some degree to a post-war "baby-boom" similar to that experienced in a number of other countries (Arnold and Phananiramai, 1975). Declining levels of mortality during the mid-1960's may also have contributed to the observed discrepancies between the enumerated and stable age distributions under age 10, although to a lesser extent than the factors cited above.

Even if allowance is made for the exaggerating effects of cyclical changes in fertility in Thailand, the fact that the age pattern of irregularities in the 1970 Thai census resembles a common pattern of age errors in censuses in developing countries

Figure 5-21. RATIO OF THE PROPORTION OF THE POPULATION ENUMERATED AT EACH AGE IN THE 1970 CENSUS OF THAILAND TO A STABLE AGE DISTRIBUTION



Source: Derived from figure 5-20.

suggests that changing fertility is not the sole source of the discrepancies observed in figure 5-21. In particular, the following aspects of the census age distribution are suggestive of error (rather than "real" distortions): (a) the deficit of population aged 0 to 4 years and corresponding surplus of population aged 5 to 9 years, which may reflect the underenumeration and/or the exaggeration of ages of children; (b) the extreme deficit of males aged 20 to 29 years, which likely reflects the typically-encountered underenumeration of males in this age range (and/or possibly migration); (c) a surplus of females aged 30 to 39 years, which has been hypothesized to reflect the exaggeration of ages of women aged 15 to 29 years in order to make their ages consistent with perceived expectations regarding age at marriage and fertility (Ewbank, 1981); and (d) extreme fluctuations at older ages, which are thought to reflect the tendency to exaggerate ages in cultures where old age is revered.

Thus, while there is reason to suspect that the age distribution of the population enumerated in the 1970 Thai census may have been distorted by swings in fertility associated with World War II, there is also reason to suspect the presence of substantial coverage and/or content error in various segments of the population which would require further investigation using other methods described in this chapter.

3.32 <u>1960 Census of Brazil</u>.--As a second example, a variant of the procedure described above recommended by the United Nations (1983) is illustrated.

The general approach and computational procedure of this variant are very similar to that described above. The basic differences between this approach and the approach described above lies in the amount of information required about the population being studied and the manner in which the accuracy of the recorded age distribution is assessed. While the two approaches should produce approximately the same results, the latter approach will prove more useful in cases of limited demographic information.

Specifically, the procedure involves the determination of a sequence of implied birth rates, $b(x)$, for 5-year age groups in the enumerated census population calculated by finding, through interpolation, the model stable population at an <u>arbitrarily chosen level of mortality</u> that has the same proportions under age x as the population enumerated in the census.

The rationale behind this approach is that in a stable population the implied birth rate for each age group will be approximately constant. Accordingly, in an actual population which has been subject to conditions of approximate stability, significant variations of the implied $b(x)$ values may be construed as being indicative of some mixture of coverage and content error in the census. A value of $b(x)$ which is higher than predicted on the basis of a genuinely stable population, for example, would indicate that too high a proportion of the population was recorded as being under age x in the census. Lower than predicted values of $b(x)$ would suggest the opposite pattern of net census error.

An important feature of this approach is that the actual level of mortality in the population under study need not be known very precisely. Conceptually, the procedure should produce reasonable results at any arbitrarily chosen level of mortality, although the selection of a model life table which approximates the actual mortality conditions in the population being studied is likely to produce better results (see United Nations, 1983: 162, for further discussion of this point).

Computationally, three steps which are very similar to those described above are required:

(1) Step 1: Calculation of the cumulated age distribution. The cumulated age distribution, or the proportion of the population under exact age x, denoted $C(x)$, is determined by summing the proportions in each age group, the $c(x)$ values described in the first example, up to the required age x.

(2) Step 2: Calculation of birth rates at a fixed level of mortality. This step involves the identification of the stable birth rate, $b(x)$, for each value of x which is implied by the level of mortality chosen and the observed values of $C(x)$ from the census. As in the first example, the appropriate values are derived through interpolation; but in this case a one-way interpolation is performed between the printed values of $C(x)$ in the model stable populations at the chosen (fixed) level of mortality. The values of $b(x)$ chosen are those which correspond to the observed $C(x)$ values from the census. This procedure is illustrated in the example below.

(3) Step 3: Plot of the estimated birth rates $b(x)$ against age x. The estimated $b(x)$ values are then plotted against values of age, x, to assess the extent to which they fall on a straight line. As noted above, departures from a straight line are indicative either of nonstability of the population or of net census error (coverage and/or content error).

To illustrate the above approach, the example using the male age distribution of the 1960 census of Brazil presented in United Nations (1983:160-62) has been reproduced. The proportions of the enumerated population in each 5-year age group, $c(x)$, and the cumulative proportions up to each age x, $C(x)$, required in step 1 of the computational procedure are shown in columns (2) and (4) of figure 5-22. The $c(x)$ values were obtained by summing the values of $C(x)$ up to the required age.

The estimated birth rates implied by the proportion of the population under each age x, $b(x)$, shown in column (5) of the table were derived by interpolating between printed values of $C(x)$ in the Coale-Demeny model stable populations corresponding to mortality level 13 of the "West" regional model life table.

Figure 5-22. AGE DISTRIBUTION AND FITTED BIRTH-RATE ESTIMATES FOR MALES, BRAZIL, 1960

Age group (1)	Proportion in each age group $c(x)$ (2)	Age x (3)	Cumulated age distribution $C(x)$ (4)	Estimated birth rate $b(x)$ (5)
0 to 4 years.......	0.1625	5	0.1625	0.04114
5 to 9 years.......	0.1477	10	0.3102	0.04316
10 to 14 years.....	0.1227	15	0.4329	0.04337
15 to 19 years.....	0.0986	20	0.5315	0.04253
20 to 24 years.....	0.0855	25	0.6170	0.04202
25 to 29 years.....	0.0727	30	0.6897	0.04151
30 to 34 years.....	0.0644	35	0.7541	0.04137
35 to 39 years.....	0.0563	40	0.8104	0.04147
40 to 44 years.....	0.0474	45	0.8578	0.04159
45 to 49 years.....	0.0398	50	0.8976	0.04186
50 to 54 years.....	0.0317	55	0.9293	0.04203
55 to 59 years.....	0.0235	60	0.9528	0.04163
60 to 64 years.....	0.0206	65	0.9734	0.04503
65 years or older..	0.0266	w	1.0000	-

Source: United Nations (1983)

To illustrate this interpolation procedure, the value of $b(25)$, the birth rate in a stable population at mortality level 13 (expectation of life = 47.11 years) having a value of $C(25)$ identical to the value recorded in the Brazilian census, was calculated as follows. The printed values of $C(25)$ in the model stable population with a birth rate, $b(25)$, of 0.03756 is $C(25) = 0.5813$, while $C(25) = 0.6193$ in the stable population with $b(25) = 0.04230$. Since these two values of $C(25)$ "bracket" the observed $C(25)$ in the Brazilian census (which equals 0.6170), interpolation is performed between these two values of $C(25)$ to derive the interpolation factors (θ) and $(1-\theta)$ from which to interpolate between the $b(25)$ values as follows:

$$\theta = \frac{0.6170 - 0.5813}{0.6193 - 0.5813}$$
$$= 0.94$$

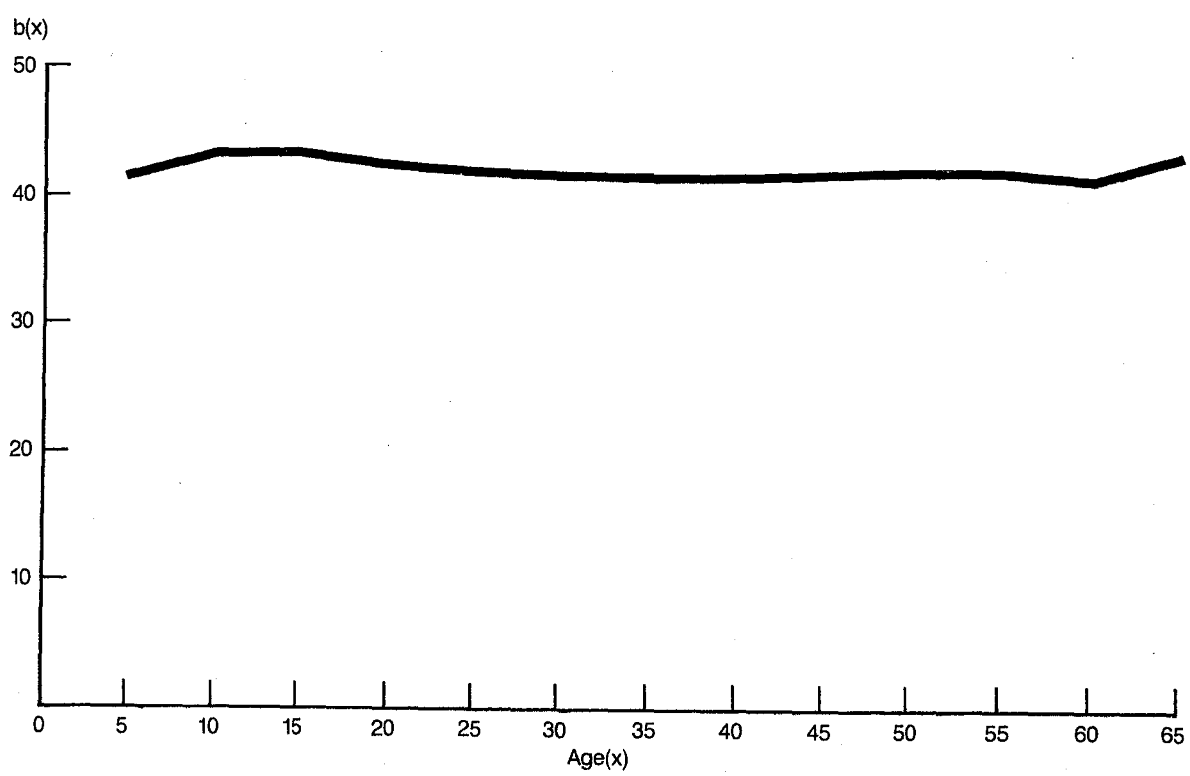
The interpolation factors (θ) and $(1-\theta)$ were then applied to the corresponding values of $b(25)$ to determine the value of $b(25)$ implied by the observed value of $C(25)$ in the Brazilian census as follows:

$$b(25) = 0.94\ (0.04230) + 0.06\ (0.03756)$$
$$= 0.04202$$

This procedure was used in estimating the values of $b(x)$ shown in column (5) of figure 5-22.

The final step in the procedure is to plot the estimated $b(x)$ values by age, x, as shown in figure 5-23. On the basis of this figure, it would appear both that the Brazilian population was approximately stable and the 1960 census male age distribution had not been seriously distorted by net census error, although some distortion at ages 10 and 15 and at the older ages is evident. The relatively modest degree of distortion in the age distribution of Brazilian males in the 1960 census is evident when compared to the plots of $b(x)$ by age for selected

Figure 5-23. PLOT OF THE ESTIMATED BIRTH-RATE, b(x), SEQUENCE BASED ON THE MODEL STABLE MALE POPULATION, WEST FAMILY, LEVEL 13, BRAZIL, 1960



Source: United Nations (1983), table 18.

African female, figure 5-24, and Latin American/Filipino male populations, figure 5-25.

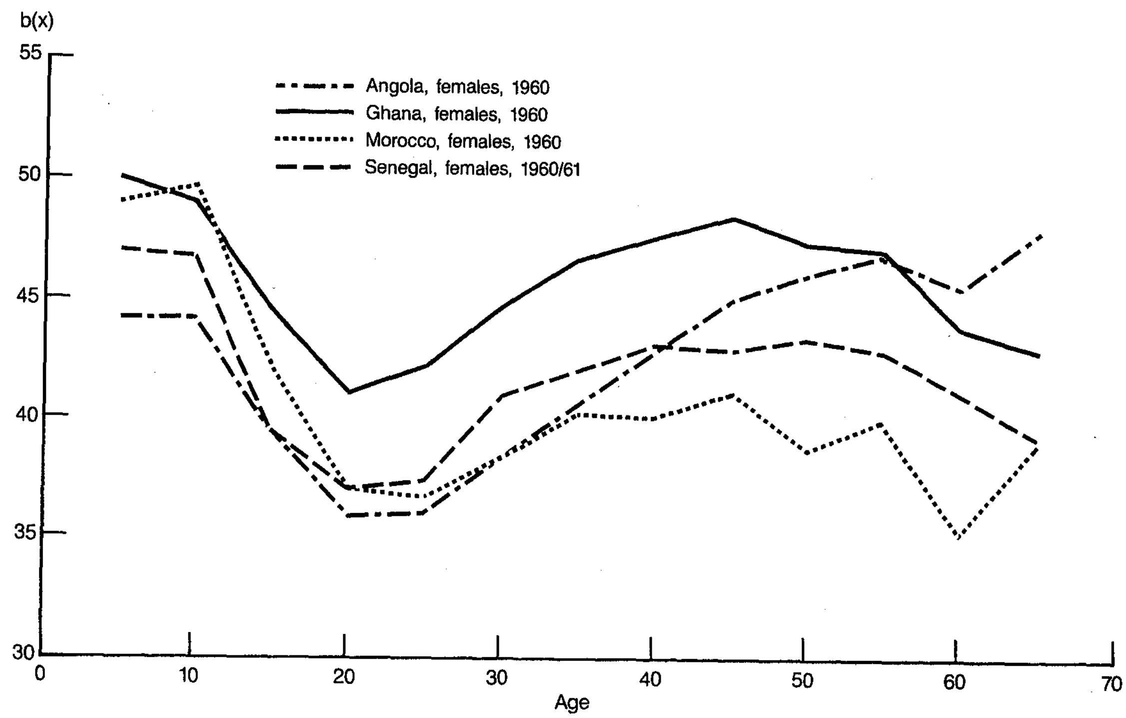
The pattern of $b(x)$ values shown in figures 5-24 and 5-25, follow frequently observed patterns of distortion in census age data. The $b(x)$ values for African females shown in figure 5-24, for example, seem to follow a pattern characteristic of populations (for both males and females) in Africa and South Asia (United Nations, 1983). This pattern of age errors is characterized by maximum values of $b(x)$ at age 10, a sharp decline up to age 20, and a gradual increase thereafter with increasing age. This pattern indicates a significant deficit of females ages 10 to 19 in census enumerations of these populations and is likely the result of age misreporting and errors made by census enumerators in estimating the ages of females in this age range. It should be noted, however, that the PES results of the 1981 censuses of India and Bangladesh indicate that omission rates tend to be relatively high in this age range (India, 1982; Bangladesh, 1983), suggesting that coverage error may also be a contributing factor to the distortion in the age distribution of at least the South Asian populations.

The pattern of $b(x)$ values for Latin American and Filipino males shown in figure 5-25 suggest a substantially different pattern of census errors. This pattern is characterized by maximum values of $b(x)$ at age 15 followed by a gradual decline with increasing age. This pattern might reflect a tendency toward increasingly greater overstatement of age among successively older cohorts of males. It would appear, however, that the distortions in the age distributions of these countries are considerably less marked than those shown in figure 5-24.

3.4 *Uses and Limitations*

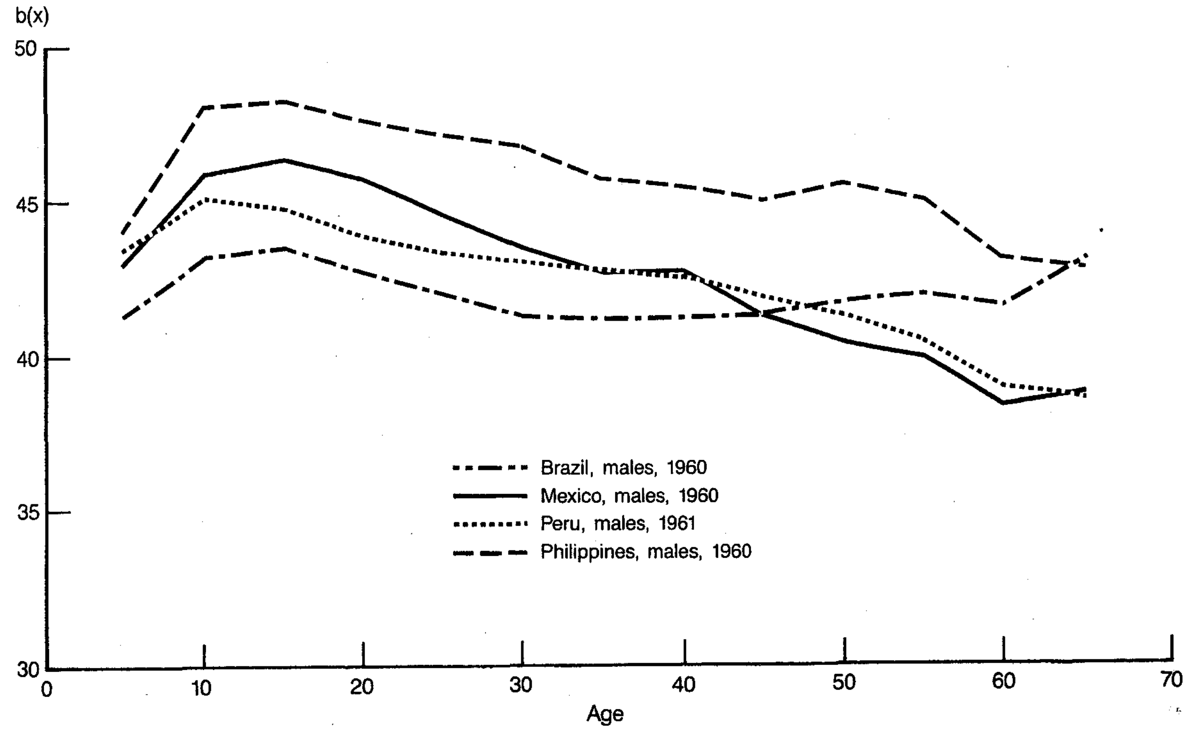
In countries in which a reasonable case can be made for approximately stable demographic conditions, stable population techniques

Figure 5-24. BIRTH RATE ESTIMATES, $b(x)$, OBTAINED FROM PROPORTIONS UNDER AGE x, $C(x)$, IN MODEL STABLE POPULATIONS FOR SELECTED AFRICAN FEMALE POPULATIONS



Source: U.N. Manual X, figure 19, panel e

Figure 5-25. BIRTH RATE ESTIMATES, b(x), OBTAINED FROM PROPORTIONS UNDER AGE x, C(x), IN MODEL STABLE POPULATIONS FOR SELECTED LATIN AMERICAN MALE POPULATIONS AND PHILIPPINE MALES



Source: U.N. Manual X, figure 19, panel f

provide a valid basis for evaluating the accuracy of census age statistics. The usefulness of the approach derives both from the robustness of the method in previous applications under a fairly wide range of recently destabilized conditions and the fact that the conditions assumed under the model (constant fertility and constant or recently declining mortality) are satisfied in a number of countries.

The recent declines in fertility experienced in a number of countries, however, limits the usefulness of the approach since the resulting estimates under either of the variants described above are sensitive to changes in levels of fertility. Substantial levels of net migration also will limit the applicability of the method, but in some cases it may be possible to overcome this limitation by adjusting the recorded census age distribution for the effects of migration prior to the application of stable population techniques. Recent declines in mortality should not significantly limit the use of the approach since changing levels of mortality have a rather small effect on the age distribution of the population. Larger mortality declines may, however, necessitate the use of a somewhat different approach as described below.

In actual application, one of the more problematic aspects of the use of stable population techniques arises in the interpretation of results. One problem is that it is often difficult to judge the extent to which discrepancies between the recorded and stable age distributions should be attributed to census error, to a departure in the population studied from the assumed condition of approximate stability, or to other factors. In some instances, distorted age distributions are the result of extraordinary events such as wars, famines, etc., and the observed discrepancies can be explained on this basis. In other cases, however, the causes of the observed discrepancies will be less apparent. Previous experience has shown that census age errors tend to follow several common patterns and the observation of one of these patterns in actual application should provide justification for the assumption of approximate stability (see United Nations, 1983 and Ewbank, 1981 for a description of typical patterns of census age errors). Failure to observe one of these common patterns should, however, cause the validity of the assumption of stability to be investigated further before accepting the observed discrepancies as estimates of census error.

Another limitation of the method is that it is often difficult to assess the relative contributions of coverage and content error to observed discrepancies in the census age distribution. One possibility for overcoming this weakness is to use stable population techniques in combination with other methods which provide estimates of net census coverage error by age and sex. Ewbank (1981), for example, illustrates the use of stable population analysis in conjunction with PES estimates of net coverage error for the 1972 census of Paraguay (see figure 5-26). In this table, the age distribution recorded in the census (column 1) was adjusted for net census undercount (column 4) on the basis of PES results and the adjusted age distribution compared with an appropriate stable age distribution. The rationale behind this approach is that in an approximately stable population and on the assumption that net census <u>coverage</u> error has been estimated accurately by the PES, the remaining discrepancies between the adjusted census and stable age distributions

Figure 5-26. STABLE POPULATION ANALYSIS OF THE REPORTED AND ADJUSTED AGE DISTRIBUTION: PARAGUAY 1972

Age	Reported by census			Adjusted for undercount		
	$c(x)$ (1)	Stable $c(x)$ (2)	$c(x)$ ÷ Stable $c(x)$ (3)	$c(x)$ (4)	Stable $c(x)$ (5)	$c(x)$ ÷ Stable $c(x)$ (6)
Males						
0 to 4 years	16.00	17.26	0.93	16.14	17.63	0.92
5 to 9 years	15.67	14.42	1.09	15.27	14.65	1.04
10 to 14 years	14.40	12.29	1.17	14.03	12.41	1.13
15 to 19 years	11.17	10.46	1.07	11.62	10.51	1.11
20 to 24 years	7.98	8.85	0.90	8.59	8.84	0.97
25 to 29 years	6.17	7.47	0.83	6.60	7.42	0.89
30 to 34 years	5.37	6.28	0.86	5.42	6.21	0.87
35 to 39 years	4.33	5.26	0.82	4.26	5.18	0.82
40 to 44 years	4.48	4.38	1.02	4.34	4.29	1.01
Females						
0 to 4 years	15.21	16.32	0.93	15.31	16.80	0.91
5 to 9 years	14.88	13.80	1.08	14.50	14.11	1.03
10 to 14 years	13.36	11.87	1.13	13.51	12.05	1.12
15 to 19 years	11.09	10.21	1.09	11.56	10.30	1.12
20 to 24 years	8.24	8.74	0.94	8.90	8.75	1.02
25 to 29 years	6.51	7.45	0.87	6.59	7.41	0.89
30 to 34 years	5.44	6.35	0.86	5.35	6.27	0.85
35 to 39 years	4.68	5.39	0.87	4.49	5.29	0.85
40 to 44 years	4.48	4.55	0.98	4.35	4.43	0.98

Source: Adapted from Ewbank (1981), based on Marks and Rumford (1978) and U.N. Demographic Yearbook (1978).

should reflect solely the effects of content (age misreporting) error. (These assumptions are never, of course, fully met in actual application).

In the case of Paraguay, the adjustment of the recorded census age distribution for net census undercount improves the "fit" between the census and stable age distributions in most age categories (compare column 3 with column 6 in figure 5-26). The fact that the adjustments are generally larger for males than for females reflects the PES estimates of higher degree of underenumeration of males in the 1972 census. However, the fact that significant discrepancies remain between the adjusted census and stable age distributions (column 6) supports the interpretation that age misreporting (content error) is a significant source of error in the census data for most age groups.

In addition to the uses described above, stable population techniques often prove useful in countries which have conducted multiple censuses in adjusting the age distributions recorded in an earlier census when stable conditions may more reasonably be assumed to have applied. The adjusted population of the earlier census may then be "projected" forward to the date of a current census based upon estimates of levels and trends in fertility, mortality, and migration in the intervening period and compared with the current census. This procedure, the cohort-component method, is described in section 5.

The procedure of accepting a stable age distribution along with the enumerated total from a previous census and projecting this population forward to the date of a subsequent census is often a useful approach in cases where changing fertility and/or mortality invalidate the direct use of stable population methods to evaluate the age distribution of the subsequent census. In a situation of declining fertility, this would in fact be a prescribed approach. In a situation of declining mortality over a significant number of years, a choice needs to be made between this approach and the use of quasi-stable population methods.

A case for the application of the cohort-component approach using a stable age distribution for an earlier census date in lieu of quasi-stable methods may be made on the basis of two points. First, the number and severity of assumptions required are considerably less under the cohort-component approach than under the quasi-stable approach. Second, and perhaps more relevant, is the fact that in most instances in which sufficient information is available for the application of quasi-stable methods (the length of time and the rate at which mortality has been declining), sufficient information also will be available to apply the cohort-component approach.

4. ANALYSES OF SUCCESSIVE CENSUSES USING ACTUAL DATA ON COMPONENTS OF POPULATION CHANGE

In countries with relatively complete systems of vital registration and/or in which evaluation studies have been undertaken so that a fairly reliable estimate of the degree of under-registration in these systems can be derived, information on the number of intercensal births, deaths, and net international migrants can be used in conjunction with the results of a previous census to evaluate the coverage of a subsequent or current census.

4.1 *Basis of method*

The basic mathematical framework for analyzing successive censuses is commonly referred to as the population balancing equation, which may be expressed symbolically as:

(5.9) $P_1 = P_0 + B - D + M$

Where:

- P_1 = the population enumerated in the census being evaluated
- P_0 = the population enumerated in a previous census
- B,D = the number of births and deaths in the period between the two censuses, respectively
- M = the number of net international migrants in the period, which is equal to the number of immigrants (I) less the number of emigrants (E); that is $M = I - E$. In the case of net immigration, M will have a positive sign; in the case of net emigration, M will have a negative sign.

The logic underlying the balancing equation is quite simple. Assuming fixed national boundaries, the population of a country can increase or decrease between any two points in time (such as between two censuses) only as a result of births, deaths, and movement of population across national boundaries. Births and immigration between the two censuses add to the population, while deaths and emigration during the intercensal period reduce it. In the absence of error in the components of the balancing equation, the equation will balance exactly.

For census evaluation purposes, the extent to which the enumerated population (P_1) falls short of (or exceeds) the population count expected on the basis of the sum of the components on the right hand side of equation (5.9) is of prime interest. The "residual" quantity needed to make the equation balance exactly, labeled "e" in equation (5.10), is referred to as the "error of closure" and represents an estimate of the relative coverage error in the two censuses. This estimate is derived by solving equation (5.10) for e.

(5.10) $P_1 = P_0 + B - D + M + e$

If a negative residual quantity e, is needed to balance equation (5.10), the enumerated total, P_1, will have fallen short of the expected total, or in other words will have been underenumerated relative to P_0. If a positive residual is required to balance the equation, a relative overenumeration of P_1 is indicated.

In actual application, there are several important limitations to this procedure. Foremost among these is that because the extent of coverage error in P_1 is estimated by a residual value in the balancing equation, this estimate is directly affected by errors in each of the other components of the equation. Underenumeration in the first census, P_0, and under-registration of births, B, or net emigration, M, will lead to a lower "expected" population and an estimate of the level of undercoverage in P_1 which is biased downward. Overenumeration of P_0, underregistration of deaths, D, and of net immigration will have the opposite effect. It is often the case that errors in the components of the balancing equation compensate for or mutually cancel other errors, as in the case of underenumeration of P_1 and underregistration of births. The implication of this is that a small residual value, e, in the balancing equation cannot necessarily be interpreted as being indicative of a low level of net census coverage error in P_1.

Accordingly, evaluation of the accuracy of the components of the balancing equation should be undertaken and appropriate corrections or adjustments made prior to using the balancing equation to evaluate the level of coverage error in P_1. In cases where births and/or deaths are known or thought to be underregistered, estimates of the degree of underregistration can be used to adjust the recorded numbers of intercensal births and deaths.

The procedure proposed by Chandrasekar and Deming (1949), which is of common theoretical origin as the PES model for evaluating census coverage outlined in chapter 2, has been utilized by a number of countries for this purpose. Other procedures for adjusting the registered number of births and deaths have been suggested by Brass (1975), Preston et al. (1980), Preston and Hill (1980), and Luther and Retherford (1985). Several of these methods are documented in United Nations (1983).

Deriving similar adjustments for the migration component of the balancing equation is typically more difficult. This is due to the less comprehensive nature of the information collected on migration by many countries and the generally lower quality of these data in comparison with birth and death registration data. As a result, it is generally the case that there is less confidence in the accuracy of migration data than in fertility and mortality data.

In countries where the volume of international migration is small or in which the volume of in- and out-migration is thought to be more or less equivalent (that is, net international migration is close to zero), the migration component of the balancing equation can be dropped without a significant adverse effect on the resulting estimate of census coverage. In countries with significant levels of net migration, however, an estimate of the number of net international migrants during the intercensal period will be required. One or more of the indirect methods for estimating international migration outlined in chapter 4 might be used in deriving such an estimate.

It should be noted, however, that the census being evaluated should not be used in deriving a "residual" estimate of the volume of net intercensal migration. The reason for this is that such residual estimates of migration also will include at least some portion of census coverage error, which when inserted into the balancing equation, will produce an estimate of census coverage error which is biased downward. The procedure used to estimate the migration component of the balancing equation must be independent of the census being evaluated.

Finally, if the population balancing equation is to be used to estimate the degree of census coverage error on an absolute basis, an adjustment for net coverage error in the previous census will be required. This adjustment is necessary because the previous census, like the census being evaluated, is subject to error. If the enumerated population from the previous census, P_0, is adjusted to account for net coverage error, the population balancing equation will yield an estimate of net coverage error in the second or current census, P_1. If such an adjustment is not made, the balancing equation will produce an estimate of the difference in coverage levels between the two censuses, or relative coverage. If the residual term, e, in the balancing equation is negative, the implication would be that the second census, P_1, was enumerated less completely than the first census, P_0. A positive residual term would imply that the second census was enumerated more completely than the first census.

4.2 *Data required*

The following data are required for the application of this method:

(1) The population counts from the census under evaluation, P_1, and from a previous census, P_0. (If an estimate of net coverage error in the current census is sought, the enumerated population from the previous census should be adjusted for net coverage error. Otherwise, the resulting estimate

will represent the relative coverage error in the second census in comparison with the first.)

(2) The number of births, deaths, and net international migrants during the intercensal period, adjusted for under-registration (to the extent possible).

4.3 *Computational procedure*

For the procedure involving analyses of successive censuses using the population balancing equation, the computational steps described below are required.

4.31 Step 1: Adjustment of registered numbers of intercensal births, deaths, and migrants.--The initial step in the procedure is to correct (to the extent possible) the numbers of births, deaths, and (where applicable) migrants recorded in the vital and immigration record systems. Where evaluation studies have been undertaken to assess the level of undercoverage in these systems, the resulting estimates can be used to inflate the registered numbers of events to reflect the true number of events more accurately. In the case of migration, it is frequently found that indirect estimates must be used due to severely deficient immigration statistics and/or the lack of suitable information with which to adjust the statistics.

4.32 Step 2: Computation of the "expected" census population.--The population "expected" on the reference date of the census being evaluated $E(P_1)$ consists of the number of persons enumerated in a previous census, P_0, preferably adjusted for net coverage error, plus the adjusted number of births, B, during the intercensal period, minus the adjusted number of intercensal deaths, D, plus or minus the number of net migrants during the intercensal period, M.

(5.11) $$E(P_1) = P_0 + B - D \pm M$$

4.33 Step 3: Calculation of the residual error or error of closure.--The residual error, e, which represents an estimate of net coverage error, is calculated by subtracting the expected census count $E(P_1)$, from the enumerated census count, P_1.

(5.12) $$e = P_1 - E(P_1)$$

As noted previously, if the enumerated population in the previous census, P_0, has been adjusted for net coverage error, the estimated residual error, e, will represent an estimate of net coverage error in P_1. If such an adjustment is not made, this estimate will represent an estimate of the relative level of net coverage error in P_1 in comparison with P_0.

4.4 *Example using population balancing equation*

To illustrate the use of the population balancing equation, the method was applied to the 1981 census of Sri Lanka. A total of 14,848,364 persons were enumerated in the 1981 census. Data from the 1971 census, in which 12,689,897 persons were enumerated, and registration data on the number of births and deaths during the intercensal period as reported by the Sri Lanka Registrar-General's Department and net international migrants as reported by the Department of Immigration and Emigration are used to assess the accuracy of the total population enumerated in 1981.

Figure 5-27 shows, for each year from 1971 to 1981, the registered number of births (column 1), deaths (column 4), and net migrants (column 8). The remainder of the data displayed in this table documents the adjustments made to the registered figures to minimize the effects of several deficiencies.

Two adjustments were made to the birth registration data shown in column (1). First,

Figure 5-27. REGISTERED AND ADJUSTED NUMBERS OF BIRTHS, DEATHS, AND NET MIGRANTS IN SRI LANKA DURING THE 1971 - 1981 INTERCENSAL PERIOD

Year	Births			Deaths				Net Migrants		
	Registered (1)	Adjusted for under-registration[1] (2)	Adjusted to census date[2] (3)	Registered (4)	Under registration adjustment factors (5)	Adjusted for under-registration[3] (6)	Adjusted to census date[2] (7)	Registered (8)	Adjusted to census date[2] (9)	Net population increase (decrease) (10)
Total	4,297,586	4,353,455	3,716,878	1,113,036	(X)	1,153,251	1,002,108	-495,000	-446,911	2,267,859
1971	382,668	387,643	88,150	97,209	.955	101,790	23,147	-11,000	-2,501	62,502
1972	385,462	390,473	390,473	103,918	.958	108,474	108,474	-10,000	-10,000	271,999
1973	367,158	371,931	371,931	100,678	.960	104,873	104,873	-50,000	-50,000	217,058
1974	365,902	370,659	370,659	119,518	.962	124,239	124,239	-54,000	-54,000	192,420
1975	375,857	380,743	380,743	115,108	.964	119,407	119,407	-31,000	-31,000	230,336
1976	380,702	385,651	385,651	106,506	.966	110,255	110,255	-52,000	-52,000	223,396
1977	389,522	394,586	394,586	103,284	.967	106,809	106,809	-52,000	-52,000	235,777
1978	404,831	410,094	410,094	93,971	.969	96,977	96,977	-40,000	-40,000	273,117
1979	417,986	423,420	423,420	94,244	.971	97,059	97,059	-44,000	-44,000	282,361
1980	407,243	412,537	412,537	89,325	.973	91,804	91,804	-101,000	-101,000	219,733
1981	420,255	425,718	88,634	89,275	.975	91,564	19,064	-50,000	-10,410	59,160

Source: Sri Lanka (1982), table 45; United Nations (1984), table 1.

[1]*Adjustment assumes constant under-registration of births of 1.3 percent.*

[2]*1971 figure is an estimate of the number of births, deaths and net migrants between the census date (October 9) and the end of the year derived by allocating the proportion (83 ÷ 365 = .2274) of events registered in 1971 to this period. The proportion (76 ÷ 365 = .2082) was used to allocate 1981 events to the period up to the census date (March 17)*

[3]*Column (6) = Column (4) ÷ Column (5)*

(X) Not applicable

the number of births registered in each year was corrected to compensate for under-registration. The completeness of birth registration in Sri Lanka is quite high, having been estimated at 98.7 percent complete in 1967 (Sri Lanka, 1970). In view of this already low level of under-registration, it was assumed for this example that no further significant improvements took place in the 1970's and the number of births registered in each year from 1971 to 1981 was revised upward by 1.3 percent as shown in column (2).

The second adjustment was necessitated by the fact that the 1971 census was conducted on October 9, 1971 while the 1981 census was taken in March 17, 1981. As a result, some portion of the registered births, deaths, and migration shown for the years 1971 and 1981 would have occurred outside of the intercensal period and accordingly should not be included in the data used to evaluate census coverage.

To compensate for this, estimates of the number of 1971 births occurring after the census date and of the 1981 births occurring prior to the census date (that is, those occurring during the intercensal period) were derived and substituted for registered numbers of births in 1971 and 1981, as shown in column (3). The adjustment factor for 1971 was derived by dividing the number of days between the census date and the end of the calendar year (83 days) by 365 days. The resulting proportion (.2274) was then multiplied by the number of registered births in 1971 to yield the estimate of 88,150 births occurring between October 9 and December 31, 1971 shown in column (3). Similarly, the number of days between the beginning of calendar year 1981 and the 1981 census date (76 days) was divided by the total number of days in the calendar year to derive an allocation factor for registered births in 1981. The resulting factor (.2082) was then multiplied by the number of registered births in 1981 and used as an estimate of the number of 1981 births which occurred during the intercensal period (88,634) as shown in column (3).

This adjustment procedure implicitly assumes that births are distributed evenly across months. While this assumption is unlikely to hold exactly due to seasonal variability in birth rates, the resulting error is not likely to be significant. An alternative to this adjustment procedure would have been to "move" the census populations to the end or beginning of the calendar years in which they were enumerated so as to coincide exactly with reference periods for the vital registration data.

Similar adjustments were made to the registered number of deaths shown in column (4). First, an adjustment for under-registration was introduced. The adjustment factors used were based on official estimates of under-registration, which showed 11 percent under-registration in 1953 (Sri Lanka, 1953) and 5 percent in 1967 (Sri Lanka, 1970). Based upon the assumption that the completeness of death registration would continue to improve to the point of becoming comparable to that of birth registration, the adjustment factors shown in column (5) were derived by extrapolating to 1981 the continuous rate of improvement in under-registration observed from 1953 to 1967. The registered number of deaths in each year (column 4) was then divided by these adjustment factors (column 5) to yield the adjusted deaths shown in column (6). The estimated number of deaths in 1971 and 1981 were then allocated to the intercensal period following the procedure described above in connection with births. The resulting estimates are shown in column (7).

The registered number of net migrants shown in column (8) were derived from data collected on the movement of passengers into and out of Sri Lanka by the Department of Immigration and Emigration. Illustrative data on the movement of passengers by nationality and year for the years 1976 to 1981 are shown in figure 5-28. These and comparable data for the years 1971 to 1975 were used to derive the estimates of the number of net migrants for each year shown in column (8).

While subject to several important limitations (see U.S. Bureau of the Census, 1975), passenger statistics provide at least an approximation of the level of net migration for Sri Lanka in the 1971-1981 period. In the absence of additional information with which to adjust the implied estimates of net migration, these data would have to be accepted for analytic purposes. For this example, the estimates of net migration for the years 1971 to 1980 were accepted as shown in column (9). An adjustment was made, however, to the data for 1971 and 1981 to allocate migrants for these years with respect to the intercensal period using the same adjustment factors that were applied to births and deaths. The "final" estimates are shown in column (9).

Column (10) shows the resulting estimates of net population change for each year during the intercensal period derived on the basis of the adjusted figures in columns (3), (7), and (9). The sum of column (10) represents the estimate of total population change for this period. This figure (2,300,859) may also be derived on the basis of the totals of columns (3), (7), and (9).

Having adjusted the vital registration and migration data, the next step in the

Figure 5-28. MOVEMENT OF PASSENGERS INTO AND OUT OF SRI LANKA, 1976 TO 1981

Type of Movement and Nationality	1976	1977	1978	1979	1980	1981
Inward						
Total...........	186,928	224,925	314,556	374,728	382,511	522,082
Sri Lanka Citizens....	46,442	52,992	102,142	100,603	98,736	142,427
Other Indians.........	21,889	20,351	25,591	39,009	47,668	61,288
Pakistanis............	2,530	1,445	1,757	2,748	2,570	3,705
Commonwealth Citizens............	19,278	25,096	29,684	36,859	44,073	59,440
Other Aliens..........	96,789	125,041	155,392	195,509	189,464	255,222
Outward						
Total...........	239,112	276,993	355,084	418,307	483,596	572,344
Sri Lanka Citizens....	53,322	66,900	117,075	122,197	137,797	185,035
Other Indians.........	24,388	22,784	27,107	41,350	48,094	53,751
Estate Residents of Indian Origin.......	44,249	38,148	28,112	22,360	17,831	24,058
Pakistanis............	2,510	1,415	1,855	2,165	1,228	1,295
Commonwealth Citizens............	19,278	24,002	28,405	28,114	6,001	5,000
Other Aliens..........	95,365	123,744	152,530	202,121	272,645	303,205

Source: Sri Lanka (1982), table 44

computational procedure is to insert these data into the balancing equation. Based upon the unadjusted 1971 census count of 12,689,897 and the adjusted data from figure 5-27, the "expected" population at the time of the 1981 census is derived as follows:

$$E(P_1) = P_0 + B - D + M$$
$$= 12{,}689{,}897 + 3{,}716{,}878 - 1{,}002{,}108 + (-446{,}911)$$
$$= 14{,}957{,}756$$

The residual difference between the expected and enumerated 1981 census population is calculated as:

$$e = E(P_1) - P_1$$
$$= 14{,}957{,}756 - 14{,}848{,}364$$
$$= 109{,}392$$

The result suggests that the Sri Lankan 1981 census was enumerated less completely than the 1971 census by an amount equivalent to approximately three-fourths of 1 percent of the expected 1981 population count $(109{,}392 \div 14{,}957{,}756 = .0073)$.

In order to derive an estimate of the level of net underenumeration in the 1981 census on an absolute basis, the U.S. Bureau of the Census (1977) estimate of undercoverage in the 1971 census of approximately 1 percent was used to inflate the 1971 census count. The expected 1981 population count and residual error were then re-calculated as follows:

$$E(P_1) = P_0 \text{ (adjusted)} + B - D + M$$
$$E(P_1) = 12{,}816{,}796 + 3{,}716{,}878 - 1{,}002{,}108 + (-413{,}911) = 15{,}117{,}655$$
$$e = E(P_1) - P_1$$
$$= 15{,}117{,}655 - 14{,}848{,}364$$
$$= 269{,}291$$

On the basis of this estimate, the estimated level of undercoverage in the 1981 census would be approximately 1.8 percent $(269{,}291 \div 15{,}117{,}655 = .01781)$.

4.5 *Uses and limitations*

The population balancing equation provides a basis for evaluating net census coverage error in instances in which two censuses have been taken and fairly reliable information on the numbers of births, deaths, and net international migrants between two censuses is available or can be derived. Typically, sufficient information to permit the adjustment of the number of recorded events for under-registration and other biases must also be available, since these often tend to be incomplete and/or deficient in other ways. Under these conditions, this method offers the advantages of being quite simple computationally and providing reasonably accurate estimates.

In many countries, however, the applicability of the method is limited by incomplete and defective data on the components of population change. Because the estimate of net census error in the balancing equation is derived as a residual of the difference between the enumerated and an expected census population, the method is sensitive to errors in any of the equation components. Accordingly, the inability to adequately correct for deficiencies in any of the statistical systems providing the data for use in the balancing equation will bias the resulting estimate of net census coverage error. The use of the balancing equation should be approached with particular caution in countries with substantial levels of international migration, since data on migration generally tend to be of lower quality than does information on the other components of population change.

In some instances, a restricted form of the balancing equation can be used to obtain estimates of net census coverage error for certain segments of the population. For example, in cases where births and deaths of infants and young children are severely under-registered, only persons 5 years and above at the time of the first census and $5 + x$ years (where x represents the length of the intercensal period) might be considered along with deaths to this population and migration in the balancing equation. Similarly, in cases where birth registration alone is severely deficient, the balancing equation might be restricted to persons aged x years and above at the time of the second census. These restricted equations, while not yielding estimates of census coverage error for the entire population, will at least provide some information on the level of census coverage for a sizable proportion of it.

Finally, it should be noted that the population balancing equation is generally not useful for obtaining estimates of net census coverage error for sub-national populations (regions or provinces, for example). The reason for this is that, in addition to the components of population change considered above, internal migration also has to be considered. In many instances, sufficiently detailed information on levels and trends of migration between sub-national areas will not be available. For most practical purposes, the use of the population balancing equation is limited to estimation of net coverage error at the national level.

5. ANALYSES OF SUCCESSIVE CENSUSES USING ESTIMATES OF THE COMPONENTS OF POPULATION CHANGE

In cases where vital registration data are not available or are deficient to such an extent that it is not possible to adjust the data in any satisfactory manner, the evaluation procedure based upon the use of the population balancing equation described in the preceding section will not be applicable. If, however, estimates of fertility and mortality levels during the intercensal period can be derived (from a demographic survey or a current or previous census, for example) these can be used in conjunction with the results of a previous census to derive an "expected" population to evaluate the results of a second census.

5.1 *Basis of method*

Essentially, this approach, the "cohort-component" method, entails the "projection" of the population enumerated in the first census to the reference date of the second census based upon estimated levels and age schedules of fertility, mortality, and migration during the intercensal period. The "expected" population is then compared with the population enumerated in the second census. This approach differs from the method based upon the use of the population balancing equation, in that intercensal births, deaths, and migration are estimated on the basis of estimates and/or assumptions regarding levels and age schedules of these parameters rather than directly available data based on registration systems. Like the population balancing equation approach, however, the resulting estimates of census error are "residual" estimates, and the cautions noted in connection with the balancing equation are also applicable. As was also the case in using the population balancing equation, the resulting estimates under the cohort-component approach are estimates of relative or differential error in the two censuses. The method may be used in conjunction with an "adjusted" population from a previous census to derive estimates of census error in the second census on an absolute basis.

The derivation of an expected population at the reference date of the second census entails three primary operations. First, the population enumerated in the first census is "survived" on a cohort-by-cohort basis to the date of the second census based upon age-specific survival rates from a life table assumed to represent intercensal mortality conditions in the population under study (see section 3 of chapter 4 for a summary of techniques of indirect estimation which may be used to select an appropriate life table). Second, the "surviving" cohort populations are adjusted to take into account intercensal immigration or emigration. Finally, the number and timing of births during the intercensal period are estimated on the basis of an assumed schedule of fertility rates and the projected population of females of child-bearing age during the intercensal period. These births are then "survived" to the date of the second census to yield an estimate of the number of children under a certain age expected in the second census. This projected population is then compared with the population enumerated in the second census.

5.2 *Data required*

The following data are required to apply the cohort-component method:

(1) The population enumerated in two successive censuses by age and sex. Age data either in single years or in 5-year age groups may be used.

(2) Life table survival rates for males and females assumed to be representative of mortality conditions during the intercensal period.

(3) A schedule of age-specific fertility rates for women aged 15 to 49 (in 5-year age groups) assumed to represent the level and age structure of fertility during the intercensal period, or other fertility estimates from which age-specific fertility schedules may be derived.

(4) An estimate of the sex ratio at birth.

(5) Estimates of the level and age pattern of net international migration during the intercensal period (where the level of net migration is substantial). Information or assumptions regarding the distribution of net migrants by time of migration will also be needed.

5.3 *Computational procedure*

The computational procedure presented below is described in a general form applicable to intercensal periods of any length. In the most general application, a series of single-year projection intervals are used to project the population to the appropriate reference date. The use of single-year projection intervals provides maximum flexibility in cases where the intercensal period is not a convenient multiple of 5, fertility and/or mortality levels have been changing rapidly and/or in a non-linear fashion during the intercensal period, or when annual migration data are to be used. Single-year projection intervals also are used in many of the widely-available computer software packages for performing population projections. In cases where the length of the intercensal period is a convenient multiple of 5 years, the procedure may be applied using grouped age data (i.e., 5-year age categories), abridged life tables, and projection intervals of 5 years. In these instances, the initial census population classified into 5-year age groups is projected over one or more 5-year projection intervals until the reference date of the second census is reached.

Accordingly, the formulae presented below describing the required computations are expressed in terms of projection periods of variable length i (where i is an index of the number of years in the projection period). Intercensal periods are understood to consist of a variable number of projection periods, depending upon the number of years between censuses and the length of the projection periods chosen. Further, it is understood that the use of single-year projection periods implies the availability of data by single years which are either directly available from the censuses or derived by "splitting" grouped data using one of several well-known methods (see U.S. Bureau of the Census, 1975, for a description of techniques for splitting age data and deriving single-year of age, unabridged, life-tables).

5.31 Step 1: Surviving the population enumerated in the first census forward to the reference date of the second census.--The initial step in the procedure involves the projection of the population enumerated in the first census to the date of the second census taking intercensal mortality into account. Expected cohort survival rates for the intercensal period, (or more precisely for each projection period), are derived from the life table(s) chosen to reflect mortality conditions in the country using the following general formula:

(5.13) $${}_nS_x = \frac{{}_nL_{x+i}}{{}_nL_x}$$

Where:

i = the length of the projection period

${}_nS_x$ = the life table survival rate for the cohort aged x to $x+n$ at the time of the first census

${}_nL_x$ = the number of life table person-years lived in the age interval x to $x+n$

${}_nL_{x+i}$ = the number of life table person-years lived in the age interval $x+i$ to $x+n+i$

For the oldest (open-ended) age category (75 years and older, for example), the following formula is applicable:

(5.14) $${}_wS_x = T_{x+i}/{}_wT_x$$

Where:

w = the oldest age attainable in the population

i = the length of the projection period

${}_wS_x$ = the life table survival rate for the population aged x and above

${}_wT_x$ = the number of life table person-years lived at ages x and above

${}_wT_{x+i}$ = the number of life table person-years lived at ages $x+i$ and above

The computation of survival rates for children born during the intercensal period also requires a different formula which is given below in connection with the estimation of the number of intercensal births (see step 6 below).

The cohort survival rates obtained in this fashion are then multiplied by the cohort population at the beginning of the projection period $({}_nP^0_x)$ to derive a preliminary expected cohort population at the end of the projection period $({}_nP_{x+i})$ as follows:

(5.15) $${}_nP^1_{x+i} = {}_nP^0_x \times {}_nS_x$$

5.32 Step 2: Adjusting for migration.--In cases where net international migration is substantial, the "survived" cohort populations calculated as described above must be adjusted to reflect the effects of migration. Ideally, the adjustments should take into account not only the direct effects of the movement of the population across national boundaries, but also the indirect effects of subsequent migrant fertility and mortality.

What is required to make appropriate adjustments are (a) actual registration data on or estimates of the number of net migrants by age and sex during the intercensal period and (b) information on or assumptions about the distribution of net migrants by time of migration. Unless actual data by single calendar years are available or the timing of migration can be determined on some other basis, it will be necessary to assume that the net migration was distributed evenly across the years in the intercensal period; that is, that the population grew (or declined) by an equal amount in each year during the intercensal period due to net migration.

The basic procedure involved is to add/subtract the number of migrants to/from the population at appropriate points in the intercensal period and treat the migrant population in the same manner with respect to fertility and mortality as the population enumerated in the first census (assuming that migrants are subject to the same schedules of fertility and mortality as the non-migrant population, an assumption which is often necessary due to lack of information on actual fertility and mortality levels of migrants).

If estimates of annual net migration are available and 1-year projection intervals are to be employed, the appropriate number of net migrants in each cohort may simply be added to or subtracted from the "survived" cohort on a year-by-year basis. In the most precise application, migrants would be added/subtracted at the mid-point of each year, although in most cases adjusting cohort populations at the beginning or end of each year will suffice.

If annual migration data are unavailable and/or 5-year projection intervals are to be used, an assumption about the distribution of migrants with regard to time of migration within the projection period will, as noted above, be required.

Under the assumption of an equal distribution of net migrants across years of the intercensal or projection period, migrants would have immigrated to or emigrated from the country <u>on average</u> at the mid-point of the period and accordingly would have been exposed to the applicable schedules of fertility and mortality over a period equal to one-half of the intercensal or projection period. On this basis, net migrants would be appropriately added to or subtracted from the "survived" population derived as described above at the mid-point of the projection period and included in or eliminated from the base population in all subsequent computations.

An additional complication to be resolved in dealing with migration concerns the allocation of migrants to census cohorts. This problem arises because migration figures usually relate to the age of the migrants at the time of migration and not at the time of a census. Accordingly, the "census" ages of net intercensal migrants classified by age at the time of migration must be estimated.

If it can be assumed that, in addition to an even distribution across months/years of the projection period, migrants in each age group (at the time of migration) are evenly distributed with respect to age, net migrants may be assigned to projection-period cohorts by assuming that one-half of the migrants would have been in the same age group at the time of migration as at the beginning of the projection period, while one-half would have "aged" into the next higher age group between the beginning of the projection period and the time of migration. Accordingly, one-half of the net projection period migrants in a particular age group would be allocated to that age group at the beginning of the projection period, while the other half would be allocated to the next younger age group at the beginning of the projection interval.

Computationally, the introduction of net migrants (classified by age) at the mid-point of the projection period, the survival of these migrants to the end of the projection period, and the allocation to cohorts may be accomplished using the following formula:

$$(5.16)\quad {}_n\hat{M}_{x+i} = \frac{1}{4}\,{}_nM_x\,(1 + {}_nS_x) + \frac{1}{4}\,{}_nM_{x+i}\,(1 + {}_nS_{x+i})$$

Where:

i = the length of the projection period

${}_{n}\hat{M}_{x+i}$ = the expected number of surviving net projection-period migrants ages $x+i$ to $x+n+i$ years at the end of the projection period

${}_{n}M_{x}$ = the estimated number of net migrants aged x to $x+n$ years at the time of migration

${}_{n}S_{x}$ = the life-table survival rate for the population aged x to $x+n$ years

${}_{n}M_{x+i}$ = the estimated number of net migrants aged $x+i$ to $x+n+i$ years at the time of migration

${}_{n}S_{x+i}$ = the life-table survival rate for the population aged $x+i$ to $x+n+i$ years

The analogous expression for the oldest (open-ended) age category is:

$$(5.17) \qquad {}_{w}\hat{M}_{x} = \frac{1}{4}\,{}_{n}M_{x-i}\,(1 + {}_{n}S_{x-i}) + \frac{1}{2}\,{}_{w}M_{x}\,(1 + {}_{w}S_{x})$$

Where:

w = the oldest age attained in the population under study

i = the length of the projection period

${}_{w}\hat{M}_{x}$ = the expected number of surviving net projection-period migrants aged x years and above at the end of the projection period

${}_{n}M_{x-i}$ = the estimated number of net projection-period migrants aged $x-i$ to $x+n-i$ years at the time of migration

${}_{n}S_{x-i}$ = the projection-period life-table survival rate for the population aged $x-i$ to $x+n-i$ years

${}_{w}M_{x}$ = the estimated number of net projection-period migrants aged x years and above at the time of migration

${}_{w}S_{x}$ = the projection-period of life-table survival rate for the population aged x years and above

Finally, the following formula is applicable to the 0 to 4 age group:

$$(5.18) \qquad {}_{4}\hat{M}_{o} = \frac{1}{4}\,{}_{4}M_{o}\,(1 + {}_{4}S_{o})$$

Where all terms are as defined above, but specifically with regard to the population aged 0 to 4 years.

A computationally simpler but less precise procedure would consist of simply adding or subtracting the estimated number of net migrants in each cohort either to or from the cohort population enumerated in the first census or the "expected" cohort population at the time of the second census derived by surviving the enumerated first census cohort population forward to the second census date as described in section 5.31. An assumption implicit in this approach, however, is that all of the intercensal migration occurred either at the beginning or at the end of the intercensal period. To the extent that this is not the case, the resulting estimates of the effects of migration on the size of cohorts enumerated in the second census will either be overstated or understated. However, in cases where the level of migration is modest, the resulting bias may be tolerable in view of the gain in computational simplicity.

5.33 Step 3: Calculation of the average number of women of childbearing age during the intercensal period.--After having made adjustments for migration (as necessary), the next step in the computational procedure is to estimate the number of births occurring during the projection period which survived to the end of the period. The first step in this procedure is to estimate the number of women "at risk" to childbearing during the projection period. Because fertility rates vary considerably by age, an estimate of the average number of women at each age in childbearing years during each projection period is needed.

The average number of women in any age group x to $x+n$ during a projection period may be calculated as the average of the population aged x to $x+n$ at the beginning of the projection period and the projected population of females aged x to $x+n$ at the end of the period. The average number of females in each age group beginning with women in the 15 to 19 age category up to and including the 45 to 49 age category would be calculated as follows:

$$ {}_n\bar{P}_x = \frac{{}_nP^0_x + {}_nP^1_x}{2} \quad (5.19) $$

Where:

${}_n\bar{P}_x$ = average number of females aged x to $x+n$ in the projection period

${}_nP^0_x$ = the number of females aged x to $x+n$ at the beginning of the projection period

${}_nP^1_x$ = the projected number of females aged x to $x+n$ at the end of the projection period

5.34 Step 4: Estimation of the number of births during the projection period.--The number of births in the projection period may be estimated by multiplying the average number of women in each age category, ${}_n\bar{P}_x$, by the assumed age specific fertility rate, denoted ${}_nf_x$, as in equation (5.20). This will yield an estimate of the number of births per year for women aged x to $x+n$; summing across ages up to age 49 will yield an estimate of the total number of births in the projection period. If 5-year grouped age data is used, this sum must be multiplied by 5 to derive an estimate for the 5-year period. Symbolically,

$$ B = \sum_{x=15}^{49} ({}_n\bar{P}_x \cdot {}_nf_x) \text{ for a 1-year projection period, or} \quad (5.20) $$

$$ B = 5 \cdot \sum_{x=15-19}^{45-49} ({}_n\bar{P}_x \cdot {}_nf_x) \text{ for a 5-year projection period} $$

Where:

B = the estimated number of births during the projection period

${}_n\bar{P}_x$ = the average number of women in the age group x to $x+n$ years during the projection period

${}_nf_x$ = the age specific fertility rate (per woman) for women aged x to $x+n$ years during the projection period

5.35 Step 5: Distributing the estimated number of births by sex.--To obtain an estimate of the number of male and female births, the estimated or known sex ratio at birth, SRB, is used as in the following equation to estimate the proportion of births which are females:

$$ PCT\ F = 1 - \frac{SRB}{1 + SRB} \quad (5.21) $$

This estimated proportion is then applied to the estimate of total births during each projection interval segment to estimate the number of female births and, by subtraction of this result from the estimated total number of births, the number of male births. For example, with a sex ratio at birth of 1.05, the proportion of female births would be estimated as $1 - 1.05/2.05 = .488$. The total number of female births would be estimated as $B^f = B \times (.488)$ and the number of male births as $B^m = B - B^f$

5.36 Step 6: Surviving intercensal births to the end of the projection period.--The size of the cohort of children born during the projection period "expected" at the end of the projection period is estimated by applying appropriate life table survival rates to the estimated numbers of male and female births. The expected population of females aged 0 to 1 or 0 to 4 years at the end of the projection period, for example, would be estimated as:

$$ {}_1P^f_0 = B^f \cdot ({}_1L_0/\ell_0) \text{ for a 1-year projection period, and} \quad (5.22) $$

$$ {}_5P^f_0 = B^f \cdot ({}_1L_0 + {}_4L_1/5 \cdot \ell_0) \text{ for a 5-year projection period} $$

Where:

$_1P_0^f$ = the "expected" population of females aged 0 to 1 at the end of the projection period

$_5P_0^f$ = the "expected" population of females aged 0 to 4 at the end of the projection period

B^f = the estimated number of female births during the 1- or 5-year projection period

$_1L_0$, $_4L_1$ = the life-table number of person years lived in the 0 to 1 and 1 to 4 age intervals, respectively

ℓ_0 = the number of births per year in the life table population, referred to as the "radix" of the life table and ordinarily equal to 100,000

The expected number of males aged 0 to 4 would be derived in a similar fashion.

5.37 Step 7: Comparison of the expected and enumerated census populations.--The final step in the procedure is to compare the population (by age and sex) enumerated in the second census with the expected population derived as described above.

5.4 *Examples*

To illustrate the use of the cohort-component method, the method was applied to the 1980 censuses of the Philippines and Indonesia. The Philippines example illustrates the use of grouped age data, an abridged life table, adjustments for intercensal migration, and 5-year projection intervals in projecting the population enumerated in the 1970 census forward to the date of the 1980 census. In the Indondesian example, a series of nine single year projection intervals are used to derive an expected 1980 census population based on the results of the 1971 census and indirect estimates of levels of fertility and mortality during the intercensal period.

5.41 1980 census of the Philippines.--In this example, the cohort-component method is used to evaluate the 1980 census of the Philippines. The 1980 census counts of population classified by age and sex are shown in columns (2) and (4) of figure 5-29. The corresponding counts from the 1970 census, which are used in deriving an expected 1980 census population, are shown in columns (1) and (3).

In this example, the 1970 census population is projected forward 5 years to the mid-point of the intercensal period (May 1975) and then

Figure 5-29. ENUMERATED 1970 AND 1980 CENSUS POPULATION OF THE PHILIPPINES, BY AGE AND SEX

(Population in thousands)

Age	Males		Females	
	1970 Census (1)	1980 Census (2)	1970 Census (3)	1980 Census (4)
Total............	18,250	24,129	18,434	23,970
0 to 4 years..........	2,965	3,933	2,871	3,733
5 to 9 years..........	3,001	3,397	2,894	3,209
10 to 14 years........	2,547	3,036	2,478	2,914
15 to 19 years........	1,983	2,567	2,097	2,689
20 to 24 years........	1,527	2,210	1,624	2,378
25 to 29 years........	1,189	1,918	1,274	1,936
30 to 34 years........	1,008	1,521	1,064	1,478
35 to 39 years........	941	1,228	958	1,191
40 to 44 years........	732	1,046	753	1,031
45 to 49 years........	626	825	656	836
50 to 54 years........	502	683	514	704
55 to 59 years........	403	529	405	566
60 to 64 years........	311	441	302	465
65 to 69 years........	191	349	197	369
70 to 74 years........	151	216	142	224
75 years or older......	164	230	189	248
Unknown age...........	10	-	19	-

Source: 1970 census population reported in Philippines (1974), table I-7; 1980 census population reported in Philippines (1983), table 3.

over a subsequent 5-year projection interval to the reference point of the 1980 census (May 1980). While the intercensal period in this case is not exactly 10 years (the 1970 census having been conducted with a May 6 reference date, while the reference date for the 1980 census was May 1), this difference will have a negligible effect on the resulting estimates. To maximize precision, one of the census populations could have been "moved" 5 days to create an intercensal period of exactly 10 years. For the purpose of this example, however, an intercensal period of exactly 10 years is assumed.

The expected 1975 population (that is, the population at the end of the first 5-year projection period) was derived in the following manner. The 1970 census population was first survived forward based upon life-table survival rates derived from various direct and indirect estimates of mortality levels during the intercensal period. These estimates are summarized in figure 5-30. Estimates for 1970 were derived by the U.S. Bureau of the Census (1984c) based upon adjusted death registration data for the 1960-1970 period (see also United Nations 1976, and Mijares 1974). The 1975 estimates

Figure 5-30. ESTIMATED AND PROJECTED INFANT MORTALITY RATES AND LIFE EXPECTANCY AT BIRTH FOR THE PHILIPPINES, SELECTED YEARS DURING THE 1970-1980 INTERCENSAL PERIOD

Mortality measure	Reference Year		
	1970	1975	1980
Infant Mortality Rate			
Males.................	92	73	66
Females...............	69	58	51
Life Expectancy at Birth			
Males.................	56.0	59.2	60.7
Females...............	62.4	64.1	65.8

Source: U.S. Bureau of the Census (1984c), tables 9 and 10.

were derived through the application of the Brass Growth Balance technique to death registration data and the results of the 1975 census (Pagtolun-an 1983) and analyses of infant mortality rates derived from the 1978 Republic of the Philippines Fertility Survey (Philippines 1979). The 1980 estimates were derived by logistically projecting the 1975 estimates of expectation of life at birth forward based upon assumed average rates of improvement of .25 and .275 years per calendar year for males and females, respectively (U.S. Bureau of the Census 1984c).

The life tables implied by these estimates were then used to derive life tables for each of the projection periods (1970-1975 and 1975-1980). The life table for the 1970-1975 period was derived by taking the average values of the 1970 and 1975 life tables. Corresponding values for the 1975-1980 period were derived in a similar fashion based on the estimated 1975 and 1980 life tables. The ${}_nL_x$ values of these projection-period life tables and the implied 5-year survival rates (${}_5S_x$) are shown in figure 5-31.

The ${}_nS_x$ values from figure 5-31 were then applied to the 1970 census counts from figure 5-29 to derive a preliminary expected count of population aged 5 years and above in 1975 shown in figure 5-32 (column 2).

The migration data used in deriving an expected 1975 population were assembled by U.S. Bureau of the Census (1984c) based upon immigration statistics of the United States and Canada, a primary point of destination for Filipino migrants, and estimates of male labor migration to middle-eastern countries based on several sources (see U.S. Bureau of the Census, 1984c). Figure 5-33 shows the estimated number of migrants by age and sex during the 1970-1975 and 1975-1980 periods. The 1970, 1975, and 1980 data were allocated with regard to the reference dates of the

Figure 5-31. VALUES OF $_5L_x$ AND $_5S_x$ FROM LIFE TABLES FOR THE 1970-1975 AND 1975-1980 PERIODS BY AGE AND SEX, PHILIPPINES

Sex and age	1970-1975		1975-1980	
	$_5L_x$	$_5S_x$	$_5L_x$	$_5S_x$
Males				
0 to 4 years.......	451,623	0.9691	459,091	0.9747
5 to 9 years.......	437,670	0.9903	447,479	0.9914
10 to 14 years.....	433,409	0.9910	443,643	0.9920
15 to 19 years.....	429,517	0.9854	440,085	0.9862
20 to 24 years.....	423,240	0.9794	434,006	0.9826
25 to 29 years.....	414,525	0.9756	426,453	0.9795
30 to 34 years.....	404,390	0.9711	417,705	0.9752
35 to 39 years.....	392,707	0.9649	407,356	0.9687
40 to 44 years.....	378,908	0.9563	394,610	0.9603
45 to 49 years.....	362,332	0.9437	378,927	0.9486
50 to 54 years.....	341,938	0.9252	359,444	0.9318
55 to 59 years.....	316,354	0.8927	334,921	0.8970
60 to 64 years.....	282,418	0.8542	300,991	0.8625
65 to 69 years.....	241,238	0.8082	259,612	0.8173
70 to 74 years.....	194,978	0.7186	212,200	0.7189
75 years or older..	252,219	0.4445	275,758	0.4475
Females				
0 to 4 years.......	461,182	0.9740	466,868	0.9786
5 to 9 years.......	449,144	0.9923	456,859	0.9933
10 to 14 years.....	445,696	0.9940	453,773	0.9945
15 to 19 years.....	443,020	0.9917	451,298	0.9914
20 to 24 years.....	439,329	0.9887	447,435	0.9906
25 to 29 years.....	434,343	0.9853	443,246	0.9870
30 to 34 years.....	427,96[illegible]	0.9812	437,469	0.9831
35 to 39 years.....	419,894	0.9766	430,087	0.9788
40 to 44 years.....	410,065	0.9715	420,959	0.9741
45 to 49 years.....	398,388	0.9638	410,042	0.9666
50 to 54 years.....	383,952	0.9497	396,350	0.9543
55 to 59 years.....	364,895	0.9226	398,228	0.9250
60 to 64 years.....	336,637	0.8854	349,864	0.8892
65 to 69 years.....	298,068	0.8359	311,090	0.8403
70 to 74 years.....	249,162	0.7531	261,419	0.7504
75 years or older..	361,285	0.4806	378,933	0.4823

Source: Derived as described in text from U.S. Bureau of the Census (1984 c), tables 10 and 14.

projection periods assuming an equal distribution of migrants by month in these years. Accordingly, two-thirds (8 months ÷ 12 months) of the net emigration in 1970 was assumed to have occurred after the reference date of the 1970 census (May 6), which is also the reference starting point of the 1970-1975 projection period. Similarly, one-third of the 1975 net emigration was assumed to have occurred within the 1970-1975 projection interval, with the remaining two-thirds allocated to the 1975-1980 projection period, along with one-third of the 1980 net emigrants.

Net emigrants during the 1970-1975 period were survived to 1975 and allocated to census cohorts following the procedure described above in step 2 of the

Figure 5-32. WORKSHEET FOR APPLYING THE COHORT-COMPONENT METHODS TO THE 1980 CENSUS OF THE PHILIPPINES

(Population in thousands)

Age and Sex	1970 Census Population[1] (1)	Preliminary 1975 Population (2)	Adjustment for Emigration (3)	Adjusted 1975 Population (4)	Preliminary 1980 Population (5)	Adjustment for Emigration (6)	Final 1980 Population (7)
Males							
0 to 4 years......	2,967	-	- 4	3,439	-	- 2	4,020
5 to 9 years......	3,003	2,876	- 9	2,867	3,352	- 6	3,346
10 to 14 years.....	2,549	2,974	- 7	2,967	2,842	- 8	2,834
15 to 19 years.....	1,984	2,526	- 6	2,520	2,943	-15	2,928
20 to 24 years.....	1,528	1,955	- 8	1,947	2,485	-18	2,467
25 to 29 years.....	1,190	1,496	-10	1,486	1,913	-27	1,886
30 to 34 years.....	1,009	1,161	- 9	1,152	1,456	-27	1,429
35 to 39 years.....	941	980	- 8	972	1,123	-19	1,104
40 to 44 years.....	732	908	- 5	903	942	-11	931
45 to 49 years.....	626	700	- 3	697	867	- 7	860
50 to 54 years.....	502	591	- 2	589	661	- 5	656
55 to 59 years.....	403	465	- 1	464	549	- 4	545
60 to 64 years.....	311	360	- 1	359	416	- 4	412
65 to 69 years.....	191	266	- 1	265	310	- 4	306
70 to 74 years.....	151	154	- 1	153	217	- 3	214
75 years or older..	164	182	- 1	181	191	- 1	190
Females							
0 to 4 years......	2,874		- 4	3,347	-	- 2	3,896
5 to 9 years......	2,897	2,799	- 8	2,791	3,275	- 6	3,269
10 to 14 years.....	2,481	2,875	- 6	2,869	2,772	- 7	2,765
15 to 19 years.....	2,099	2,466	- 7	2,459	2,853	- 8	2,845
20 to 24 years.....	1,626	2,082	-14	2,068	2,438	-14	2,424
25 to 29 years.....	1,275	1,608	-20	1,588	2,049	-25	2,024
30 to 34 years.....	1,065	1,256	-17	1,239	1,567	-25	1,542
35 to 39 years.....	959	1,045	-11	1,034	1,218	-14	1,204
40 to 44 years.....	754	937	- 7	930	1,012	- 7	1,005
45 to 49 years.....	657	733	- 4	729	906	- 5	901
50 to 54 years.....	515	633	- 3	630	705	- 4	701
55 to 59 years.....	405	489	- 3	486	601	- 5	596
60 to 64 years.....	302	374	- 3	371	450	- 6	444
65 to 69 years.....	197	267	- 2	265	330	- 5	325
70 to 74 years.....	142	165	- 1	164	223	- 4	219
75 years or older..	189	198	- 1	197	218	- 1	217

[1] *Age unknowns from figure 5-29 distributed proportionally across age categories.*

computational procedure. The estimate of -10,000 for males aged 25 to 29 in 1975 shown in column (3) of figure 5-32, for example, was derived by applying formula (5.16) to the net emigration estimates for males aged 20 to 24 and 25 to 29 from figure 5-33 and the age-specific life table survival probabilities from figure 5-31.

$$
\begin{aligned}
{}_5M_{25} &= \frac{1}{4}\,{}_5M_{20}\,(1 + {}_5S_{20}) + \frac{1}{4}\,{}_5M_{25}\,(1 + {}_5S_{25}) \\
&= [(.25)\ (9750) + (.25)\ (9750)\ (.9794)] \\
&\quad + [(.25)(10066) + (.25)(10066)(.9756)] \\
&= 4825 + 4972 \\
&= 9797 \cong 10{,}000
\end{aligned}
$$

The estimates for the other age categories (ages 5 and above) shown in column (3) of figure 5-32 were derived in a similar fashion on the basis of formulae (5.16) and (5.17). These estimates were then used to adjust the preliminary 1975 cohort counts from column (2) as shown in column (4).

The expected 1975 population aged 0 to 4 years was derived by estimating the number of births (by sex) during the 1970-1975 period, surviving the estimated number of births to the end of the 1970-1975 projection interval, and adjusting for the estimated level of net emigration during the projection period. The number of births during this period was estimated based on the average number of females in each age group from 15 to 49 years and estimates of age-specific fertility rates during the projection period. Estimates of age-specific fertility rates (ASFR) for various reference points from 1968 to 1980 are shown in figure 5-34. For this example, the estimate for 1973 derived from the 1973 National Demographic Survey (Concepcion 1974)

Figure 5-33. ESTIMATED NUMBER OF NET EMIGRANTS FROM THE PHILIPPINES BY AGE AND SEX: 1970-1975 AND 1975-1980

Age	1970-1975		1975-1980	
	Males	Females	Males	Females
0 to 4 years.......	8,540	8,124	4,300	4,012
5 to 9 years.......	8,655	8,377	8,493	8,040
10 to 14 years.....	6,024	4,574	7,179	6,593
15 to 19 years.....	6,739	8,412	12,866	8,726
20 to 24 years.....	9,750	19,572	23,504	19,699
25 to 29 years.....	10,066	19,785	31,434	31,058
30 to 34 years.....	8,814	13,592	23,916	18,721
35 to 39 years.....	6,496	9,298	14,167	9,025
40 to 44 years.....	3,168	4,322	9,117	5,757
45 to 49 years.....	1,934	2,728	5,317	4,128
50 to 54 years.....	1,575	2,799	3,920	4,220
55 to 59 years.....	1,379	2,764	3,503	6,448
60 to 64 years.....	1,536	2,759	4,200	6,790
65 to 69 years.....	1,229	2,051	3,494	4,564
70 to 74 years.....	837	1,101	3,134	4,042
75 years or older..	379	272	1,319	1,799

Source: Based on estimates prepared for and summarized in U.S. Bureau of the Census (1984 c), table 36.

Figure 5-34. ESTIMATED AND PROJECTED AGE-SPECIFIC AND TOTAL FERTILITY RATES FOR THE PHILIPPINES FOR SELECTED YEARS, 1968 to 1980

(Rates shown are per-woman)

Age of woman	Reference Year				
	1968	1973	1975	1977	1980
Total fertility rate..	6.05	5.52	5.24	5.01	4.68
15 to 19 years............	.063	.054	.053	.041	.038
20 to 24 years............	.238	.221	.216	.204	.191
25 to 29 years............	.306	.269	.245	.236	.221
30 to 34 years............	.276	.250	.235	.241	.225
35 to 39 years............	.206	.188	.180	.172	.161
40 to 44 years............	.100	.095	.091	.089	.083
45 to 49 years............	.021	.026	.029	.018	.017

Source: U.S. Bureau of the Census (1984), tables 27 and 28.

was used as an approximation of the average level of fertility during the 1970-1975 period. Alternatively, an estimate for the exact mid-point of the projection interval could have been obtained by interpolating between the 1968 and 1975 estimates.

The estimated number of births in the 1970-1975 period was calculated as indicated in figure 5-35. The average number of women in each age category of the child bearing years during the projection period was derived as the average of the number of women in each age category at the beginning of the projection period (from column 1 of figure 5-32) and the projected number at the end of the projection interval (column 4 of figure 5-32). The average number of women "at risk" in each age category was multiplied by the appropriate age-specific fertility rate to derive the estimated number of annual births during the projection period shown in figure 5-35. The total number of projection period births was estimated by multiplying the result by 5 (5 x 1,489,000 = 7,445,000 for 1970-1975). Assuming a sex-ratio-at-birth of 105, 48.8 percent of these births were assumed to have been females (.488 x 7,445,000 = 3,633,160) and 51.2 percent to have been males (.512 x 7,445,000 = 3,811,800).

Male and female births were then survived forward to May 1975 on the basis of the life table survival probabilities shown in figure 5-31 and formula (5.18). For males, the survival probability of .9033 ($_5L_0/5 \cdot \ell_0$ = 451,623/500,000 = .9033) was used; and for females .9224 (461,182/500,000) was used. Accordingly, an estimated 3,443,235 males (3,811,840 x .9033) and 3,351,227 females (3,633,160 x .9224) were expected to have survived to the end of the projection period.

These estimates were then adjusted for net emigration during the projection period on the basis of the migration estimates shown in figure 5-33 and formula (5.18). The adjustment for males was calculated as follows:

$$
\begin{aligned}
{}_5M_0 &= \frac{1}{4} \; {}_5M_0 \; (1 + {}_5S_0) \\
&= [(.25)(8540) + (.25)(8540)(.9033)] \\
&= 2,135 + 1,929 \\
&= 4,064 \\
&\cong 4,000
\end{aligned}
$$

The corresponding estimate for females was 3,900.

These estimates were then subtracted (since they represent net emigration) from the estimated number of surviving male and female births to yield the expected counts of males and females aged 0 to 4 at the end of the 1970-1975 projection period shown in figure 5-32.

The projected 1975 population shown in column (4) of figure 5-32 was then used as the starting population for the 1975-1980

Figure 5-35. WORKSHEET SHOWING COMPUTATIONS OF THE ESTIMATED NUMBER OF ANNUAL BIRTHS IN THE PHILIPPINES: 1970-1975 AND 1975-1980

Age of Women	1970 - 1975					1975 - 1980				
	1970	1975	Avg	ASFR	Births	1975	1980	Avg	ASFR	Births
15 to 19 years.....	2,099	2,459	2,279	.054	123	2,459	2,845	2,652	.041	109
20 to 24 years.....	1,626	2,068	1,847	.221	408	2,068	2,424	2,246	.204	458
25 to 29 years.....	1,275	1,588	1,432	.269	385	1,588	2,024	1,806	.236	426
30 to 34 years.....	1,065	1,239	1,152	.250	288	1,239	1,542	1,391	.241	335
35 to 39 years.....	959	1,034	997	.188	187	1,034	1,204	1,119	.172	193
40 to 44 years.....	754	930	842	.095	80	930	1,005	968	.089	86
45 to 49 years.....	657	729	693	.026	18	729	901	815	.018	15

projection interval and the steps described above repeated using the 1975-1980 life-table survival rates from figure 5-31, the estimates of net emigration from figure 5-33, and the estimated schedule of age-specific fertility rates for 1977 from figure 5-34. The steps are documented in columns (5) through (7) of figure 5-32. The final results of this procedure, the expected population at the time of the 1980 census, are shown in column (7) of figure 5-32 and columns (2) of figure 5-36 (for males and females, respectively).

The results of the comparison between the expected and enumerated 1980 census populations are shown in figure 5-36. Differences in terms of thousands of population are shown in column (3), and in percentage terms in column (4).

Overall, the differences between the expected and enumerated male populations is small, suggesting that the level of completeness of enumeration of males in the 1980 census was very close to that in the 1970 census, while females were enumerated 1.67 percent less completely in the 1980 census than in 1970. This sex differential in relative coverage would appear to have resulted primarily from differences in coverage in the 30 to 39 ages range, where the expected population of females consistently exceeds the enumerated population while the opposite is true for males. This pattern raises the possibility that the migration data used in deriving the expected populations may have been flawed.

In several other respects, the pattern of error suggested by figure 5-36 conforms to patterns of error frequently observed in population censuses. The deficit of population aged 0 to 4, for example, is likely to reflect the tendency to undercount infants and young children. The significant excess of population aged 10 to 14 in the enumerated as compared to the expected 1980 population suggests that the population aged 0 to 4 in the 1970 census may have been seriously underenumerated. The deficit of population aged 15 to 24 of both sexes also is observed quite frequently, likely reflecting difficulties typically encountered in enumerating highly mobile segments of the population. The large excesses of population at age 60 years and above are likely to reflect severe age exaggeration.

The results, then, suggest that the 1980 census was underenumerated by approximately 0.8 percent (for both sexes combined) in comparison with the 1970 census. If the U.S. Bureau of the Census (1984c) estimate of 4.7 percent net underenumeration in the 1970 census is accepted, the implied level of net undercoverage in the 1980 census of the Philippines would be roughly 5.5 percent.

5.42 <u>1980 census of Indonesia</u>.--In this example, the cohort-component method is used to evaluate the 1980 census of Indonesia. This example illustrates the use of a series of single-year projections in a situation where the number of years separating successive censuses is not a convenient multiple of 5 (the most recent previous Indonesian census having been conducted in 1971).

The population enumerated in the 1980 census classified by age (in 5-year age groups) and sex is shown in columns (3) and (6) of figure 5-37 (for males and females, respectively). The male and female population enumerated in the 1971 census, which is used as the basis for deriving an expected population at the time of the 1980 census, is shown in columns (1) and (4), respectively.

In this example, the adjusted 1971 census population counts shown in columns (2) and (5) are used to derive an expected 1980 census population. These adjusted census counts were derived at the U.S. Bureau of the Census (1984b) using a cohort-component approach similar to

Figure 5-36. ENUMERATED AND EXPECTED 1980 CENSUS POPULATION OF THE PHILIPPINES, BY AGE AND SEX
(Population in thousands)

Age	Enumerated Population (1)	Expected Population (2)	Difference (3) = (1) - (2)	Percentage Difference $(4) = \frac{(3)}{(2)} \times 100$
Male				
All ages.....	24,129	24,128	+1	+0.00
0 to 4 years....	3,933	4,020	-87	-2.16
5 to 9 years....	3,397	3,346	+51	+1.52
10 to 14 years..	3,036	2,834	+202	+7.13
15 to 19 years..	2,567	2,928	-361	-12.33
20 to 24 years..	2,210	2,467	-257	-10.42
25 to 29 years..	1,918	1,886	+32	+1.70
30 to 34 years..	1,521	1,429	+92	+6.44
35 to 39 years..	1,228	1,104	+124	+11.23
40 to 44 years..	1,046	931	+115	+12.35
45 to 49 years..	825	860	-35	-4.07
50 to 54 years..	683	656	+27	+4.12
55 to 59 years..	529	545	-16	-2.94
60 to 64 years..	441	412	+29	+7.04
65 to 69 years..	349	306	+43	+14.05
70 to 74 years..	216	214	+2	+0.93
75 years and older....	230	190	+40	+21.05
Female				
All ages.....	23,970	24,377	-407	-1.67
0 to 4 years....	3,733	3,896	-163	-4.18
5 to 9 years....	3,209	3,269	-60	-1.84
10 to 14 years..	2,914	2,765	+149	+5.38
15 to 19 years..	2,689	2,845	-156	-5.48
20 to 24 years..	2,378	2,424	-46	-1.90
25 to 29 years..	1,936	2,024	-88	-4.35
30 to 34 years..	1,478	1,542	-64	-4.15
35 to 39 years..	1,191	1,204	-13	-1.08
40 to 44 years..	1,031	1,005	+26	+2.59
45 to 49 years..	836	901	-65	-7.21
50 to 54 years..	704	701	+3	+0.43
55 to 59 years..	566	596	-30	-5.03
60 to 64 years..	465	444	+21	+4.73
65 to 69 years..	369	325	+44	+13.54
70 to 74 years..	224	219	+5	+2.28
75 years and older....	248	217	+31	+14.29

Source: Enumerated 1980 Census Population reported in Philippines (1983); expected population derived as described in text.

that being described here. Under the assumption that the adjusted 1971 census figures accurately portray the Indonesian population at the time of the 1971 census (September 6 through 24), these data may be used to derive an estimate of net error in the 1980 census on an absolute basis.

The cohort-component procedure was applied to the adjusted 1971 census counts by single-years of age (not shown) and sex in the following manner. First, the 1971 census population was survived using estimated life-table survival probabilities corresponding to different points in the intercensal period. The mortality estimates used in developing the life tables are summarized in figure 5-38. The estimates for the 1961-1971 period were derived on the basis of analyses of adjusted census survival ratios between the 1961 and 1971 censuses and the application of Brass mortality techniques to data from the 1971 census and the 1964-1965 National Socio-Economic Survey (U.S. Bureau of the Census 1975a); the 1976 estimates were derived from analyses of sample vital registration data (Gardiner 1978), data from the 1979 National Socio-Economic Survey (Indonesia 1981), and the 1980 census (Indonesia 1983). The 1980 estimates were derived by projecting the rate of increase in life expectancy at birth in the 1960's to 1976 forward to 1980 (U.S. Bureau of the Census 1984a).

Figure 5-37. ENUMERATED AND ADJUSTED 1971 CENSUS POPULATION AND ENUMERATED 1980 CENSUS POPULATION OF INDONESIA, BY AGE AND SEX

(Population in thousands)

Age	Males			Females		
	1971 Census		1980 Census	1971 Census		1980 Census
	Enumerated (1)	Adjusted (2)	Enumerated (3)	Enumerated (4)	Adjusted (5)	Enumerated (6)
All ages.........	68,339	62,493	72,952	60,029	63,251	73,825
0 to 4 years..........	9,606	10,389	10,816	9,493	10,117	10,375
5 to 9 years..........	9,525	9,072	10,832	9,237	8,901	10,400
10 to 14 years........	7,353	8,190	9,132	6,826	8,023	8,487
15 to 19 years........	5,588	7,228	7,513	5,738	7,085	7,771
20 to 24 years........	3,602	4,771	5,979	4,429	4,813	7,023
25 to 29 years........	3,978	4,394	5,613	4,947	4,513	5,731
30 to 34 years........	3,690	4,212	4,023	4,214	4,394	4,144
35 to 39 years........	3,948	3,824	4,191	4,031	4,031	4,359
40 to 44 years........	3,064	3,050	3,644	3,038	3,250	3,776
45 to 49 years........	2,427	2,351	3,013	2,223	2,526	3,137
50 to 54 years........	1,903	1,783	2,718	1,961	1,958	2,692
55 to 59 years........	1,126	1,289	1,721	1,100	1,447	1,670
60 to 64 years........	1,082	861	1,559	1,256	991	1,669
65 to 69 years........	549	542	811	594	666	903
70 to 74 years........	510	338	689	528	470	842
75 years or older.....	380	198	688	406	338	837
Age unknown...........	7		11	8		9

Source: 1971 enumerated population reported in Indonesia (1975), table 2; 1971 adjusted population reported in U.S. Bureau of the Census (1984), table 5; 1980 census population reported in Indonesia (1983), table 2.

Figure 5-38. ESTIMATED AND PROJECTED INFANT MORTALITY RATES AND LIFE EXPECTANCY AT BIRTH FOR INDONESIA, SELECTED YEARS DURING THE 1971-1980 INTERCENSAL PERIOD

Mortality Measure	Reference Year(s) 1961-1971	1976	1980
Infant Mortality Rate			
Males.................	155	120	108
Females..............	142	104	93
Life Expectancy at Birth			
Males.................	37.4	49.2	52.0
Females..............	40.0	51.9	55.0

Source: U.S. Bureau of the Census (1984), tables 9 and 10.

The ${}_nL_x$ values from the abridged life tables derived on the basis of these mortality estimates are shown in figure 5-39. Life-table values for single years of age were derived by applying conventional "splitting" formulae to the grouped life-table data (see U.S. Bureau of the Census 1975a). Life-table values for the years between those for which estimates were available were obtained by interpolating linearly between the values for the years 1971, 1976, and 1980. The single-year ${}_nL_x$ values were then used to survive the adjusted 1971 census population yearly during the 9-year intercensal period to September 1980, the reference point of the 1980 census.

In the case of Indonesia, net international migration during the intercensal period was of a sufficiently small magnitude that it may be safely ignored in the projection procedure. Accordingly, the expected number of persons in 1971 who were alive at the time of the 1980 census is derived solely on the basis of the 1971 adjusted census counts and estimated mortality schedule as shown in column (2) of figure 5-40.

The "survived" population of females was then used to estimate the number of intercensal births and expected survivors of these births at the date of the second census. The assumed level and age-structure of fertility at various points in the intercensal period used in deriving the estimates of the number of births are summarized in figure 5-41. The estimates for the 1971-1975 period were derived from the 1976 Intercensal Population Survey (Suharto and Cho 1978), while those for 1976-1979 were based upon "own-children" estimates from the 1980 census (Indonesia 1982). Fertility estimates for 1980 were derived on the basis of a projected continuing decline in fertility levels by the U.S. Bureau of the Census (1984a).

Estimates of the number of births in each year during the projection period were obtained by applying the age-specific fertility schedules from figure 5-41 (assumed to be the same for all ages within each 5-year age group) to the average number of females in each (1-year) age category during each single-year projection interval. The expected number of births of each sex was derived assuming a sex-ratio at birth of 1.05 on the basis of formula (5.19).

Finally, the expected number of births in each year was survived forward to reference point of the 1980 census using the life-table survival probabilities summarized in figure 5-39 on the basis of formula (5.20). The resulting estimates of the number of survivors of the children born in the intercensal period (that is, those aged 0 to 8 years) at the reference point of the 1980 census are shown in column (2) of figure 5-40.

The enumerated and expected 1980 Indonesian census populations are compared in figure 5-40. Column (3) shows the differences in the expected population from that actually

Figure 5-39. VALUES OF ${}_nL_x$ FROM LIFE TABLES CORRESPONDING TO ESTIMATED AND PROJECTED LEVELS OF MORTALITY FOR INDONESIA, SELECTED YEARS DURING THE 1971-1980 INTERCENSAL PERIOD

Sex and age	Reference year(s)		
	1961-1971	1976	1980
Males			
Under 1 year.......	90,066	91,953	92,763
1 to 4 years.......	308,785	335,479	342,715
5 to 9 years.......	358,373	404,297	416,747
10 to 14 years.....	349,076	397,281	410,840
15 to 19 years.....	335,226	391,956	406,215
20 to 24 years.....	320,273	385,662	400,751
25 to 29 years.....	304,516	377,996	394,071
30 to 34 years.....	288,232	368,458	385,605
35 to 39 years.....	270,676	356,117	374,448
40 to 44 years.....	250,869	339,965	359,578
45 to 49 years.....	227,637	318,910	339,711
50 to 54 years.....	199,263	291,669	313,338
55 to 59 years.....	164,876	256,640	278,618
60 to 64 years.....	124,146	214,063	235,630
65 to 69 years.....	81,026	166,030	186,139
70 to 74 years.....	43,440	115,928	133,078
75 years or older..	23,080	111,660	134,736
Females			
Under 1 year.......	90,646	93,266	93,775
1 to 4 years.......	317,230	343,606	350,354
5 to 9 years.......	374,134	415,044	427,045
10 to 14 years.....	359,920	408,067	421,431
15 to 19 years.....	345,829	403,214	417,417
20 to 24 years.....	330,708	397,038	412,246
25 to 29 years.....	315,250	389,012	405,437
30 to 34 years.....	299,632	378,817	396,676
35 to 39 years.....	283,253	366,575	385,987
40 to 44 years.....	264,697	352,229	373,144
45 to 49 years.....	243,527	335,195	357,451
50 to 54 years.....	218,152	313,464	336,940
55 to 59 years.....	185,510	282,356	307,012
60 to 64 years.....	145,616	240,454	266,013
65 to 69 years.....	103,702	190,318	215,620
70 to 74 years.....	64,767	136,347	159,213
75 years or older..	54,856	140,809	175,456

Source: U.S. Bureau of the Census (1984), tables 10 and 14.

enumerated in terms of numbers of persons while column (4) presents these differences in terms of percentage differences.

In terms of overall coverage, the results suggest that the 1980 census count of males was underenumerated by slightly above 6 percent and the count of females by just under 6 percent. With regard to net census error by age, the results suggest an underenumeration of children aged 0 to 4 of about 10 percent (which may be accounted for partially by the transfer of population into the 5 to 9 age category due to age misreporting); a significant underenumeration of young adults aged 15 to 34, particularly among males; and significant overenumeration of population at ages 60 and above, likely the result of a widespread tendency to exaggerate ages.

5.5 *Uses and limitations*

The cohort-component method described in this section is applicable in situations where registration data are non-existent or deficient to such an extent that satisfactory adjustment is not possible, but in which two censuses and sufficient information from which to derive indirect estimates of prevailing fertility and mortality levels are available. Since this situation describes a fairly large number of developing countries, the frequent use of this approach is not surprising.

Figure 5-40. COMPARISON OF THE ENUMERATED AND EXPECTED 1980 INDONESIAN CENSUS POPULATIONS, BY AGE AND SEX

(Population in thousands)

Age	Enumerated Population[1] (1)	Expected Population (2)	Difference (3) = (1) - (2)	Percentage Difference $(4) = \frac{(3)}{(2)} \times 100$
Male				
All ages.....	72,952[2]	77,765	-4,813	-6.19
0 to 4 years....	10,816	12,023	-1,207	-10.04
5 to 9 years....	10,832	10,364	+468	+4.52
10 to 14 years..	9,132	9,429	-297	-3.15
15 to 19 years..	7,513	8,630	-1,117	-12.94
20 to 24 years..	5,979	7,824	-1,845	-23.58
25 to 29 years..	5,613	6,656	-1,043	-15.67
30 to 34 years..	4,023	4,386	-363	-8.28
35 to 39 years..	4,191	4,110	+81	+1.97
40 to 44 years..	3,644	3,841	-197	-5.13
45 to 49 years..	3,013	3,335	-322	-9.66
50 to 54 years..	2,718	2,524	+194	+7.69
55 to 59 years..	1,721	1,825	-104	+5.70
60 to 64 years..	1,559	1,261	+298	+23.63
65 to 69 years..	811	799	+12	+1.50
70 to 74 years..	689	445	+244	+54.83
75 years or older.....	688	315	+373	+118.41
Female				
All ages.....	73,825[2]	78,353	-4,528	-5.78
0 to 4 years....	10,375	11,700	-1,325	-11.33
5 to 9 years....	10,400	10,177	+223	+2.19
10 to 14 years..	8,487	9,239	-752	-8.14
15 to 19 years..	7,771	8,486	-715	-8.43
20 to 24 years..	7,023	7,679	-656	-8.54
25 to 29 years..	5,731	6,541	-810	-12.38
30 to 34 years..	4,144	4,437	-293	-6.60
35 to 39 years..	4,359	4,232	+127	+3.00
40 to 44 years..	3,776	4,045	-269	-6.65
45 to 49 years..	3,137	3,594	-457	-12.72
50 to 54 years..	2,692	2,793	-101	-3.62
55 to 59 years..	1,670	2,057	-387	-18.81
60 to 64 years..	1,669	1,451	+218	+15.02
65 to 69 years..	903	936	-33	-3.53
70 to 74 years..	842	538	+304	+56.51
75 years or older.....	837	449	+388	+86.41

Source: Enumerated 1980 census population reported in Indonesia (1983); expected population reported in U.S. Bureau of the Census (1984) derived as described in text.

[1] *The population of East Timor was excluded from these calculations.*

[2] *The figures for the total population enumerated in the 1980 census include 11,000 males and 9,000 females whose ages were not recorded.*

Figure 5-41. ESTIMATED AND PROJECTED AGE-SPECIFIC AND TOTAL FERTILITY RATES FOR INDONESIA FOR SELECTED YEARS DURING THE 1971-1980 INTERCENSAL PERIOD

(Rates shown are per woman)

Age of woman	Reference year(s)		
	1971-1975	1976-1979	1980
Total fertility rate..	5.20	4.68	4.38
15 to 19 years............	.127	.116	.105
20 to 24 years............	.265	.248	.231
25 to 29 years............	.256	.232	.223
30 to 34 years............	.199	.177	.168
35 to 39 years............	.118	.104	.096
40 to 44 years............	.057	.046	.042
45 to 49 years............	.018	.013	.012

Source: U.S. Bureau of the Census (1984), tables 27 and 28.

As is the case in the use of the population balancing equation, the limited amount of information typically available on international migration often proves to be the most problematic aspect of the application of the cohort-component approach. In cases where the volume of net migration is substantial, a disproportionate share of attention in the application of the method will have to be devoted to the development of plausible estimates of this component.

For practical purposes, the method is a direct substitute for the evaluation approach based upon the population balancing equation in that estimates of intercensal births and deaths based upon indirect estimates of fertility and mortality levels are substituted for direct data from a vital registration system. As such, the method suffers from many of the same weaknesses as the population balancing equation approach, the major weakness being the "residual" nature of the resultant estimate of net census error. This may be a particularly serious problem in countries with limited or nonexistent vital registration systems. In such cases, indirect estimates of fertility and mortality must be used. Such estimates can be quite volatile due to such factors as violation of the assumptions underlying the techniques used and errors in the survey data (age reporting errors can be particularly problematic). The availability of indirect estimates from two or more surveys and/or prior censuses are helpful to the extent that consistent estimates from multiple applications of one or more indirect techniques can be used instead of having to rely upon the results of a single application.

In general, the usefulness of the method will depend upon the quantity and quality of demographic information available about the population under study. Where information about levels of fertility, mortality, and migration between two censuses is limited and/or unreliable, the method should be used cautiously. In such a case, the preparation of a range of estimates based upon different sets of assumptions about vital rates and migration would be prudent. In cases where sufficient information exists to derive reliable estimates of demographic parameters, the method is perhaps the most powerful among the alternative demographic approaches for the evaluation of censuses, particularly since it provides age- and sex-specific estimates of net census error.

6. *ANALYSIS OF COHORT SURVIVAL RATES*

An approach which has proven useful in previous applications is based upon the comparison of the size of birth cohorts enumerated in successive censuses (see Coale and Zelnick 1963; Coale and Rives 1973; and Mukherjee 1976 for example applications).

In a population which is closed to migration, actual changes in the number of persons in any particular birth cohort between two census dates will be due solely to mortality. Accordingly, in the absence of census errors, the ratio of the number of persons in a birth cohort enumerated in the second census to the number enumerated in the first census should approximate the survival rate that would be expected on the basis of prevailing mortality conditions. In other words, the observed intercensal cohort survival rates should be comparable to life-table cohort survival rates applicable to the population under study. In populations where net intercensal migration is negligible, departures of observed intercensal cohort survival rates from those expected may be construed as evidence of errors in one or both censuses. In populations which have experienced significant net intercensal migration, the "expected" survival rates must be modified to reflect the effects of migration on the relative size of the cohorts as enumerated in the two censuses. This may be done by adjusting the size of the cohorts enumerated in one of the two censuses.

6.1 *Basis of method*

Intercensal cohort survival rates are defined by the following expression:

$$(5.23) \qquad {}_nS_x = \frac{{}_nP^1_{x+i}}{{}_nP^0_x}$$

Where:

i = the length of the intercensal period

${}_nS_x$ = the census survival rate for the cohort aged x to $x+n$ years at the time of the first census

${}_nP^1_{x+i}$ = the population aged $x+i$ to $x+i+n$ years enumerated in the second census

${}_nP^0_x$ = the population aged x to $x+n$ in the first census

In the example applications presented below, two ways of using intercensal cohort survival rates for evaluating a census are illustrated. In the first application, discrepancies between the cohort survival rates implied by successive censuses and those expected on the basis of a life table are analyzed graphically. In the second application, the intercensal cohort and "expected" life-table survival rates are combined into a ratio measure of the following form:

$$(5.24) \qquad {}_nR_x = \frac{{}_nP^1_{x+i} \div {}_nP^0_x}{{}_nL_{x+i} \div {}_nL_x}$$

Where:

i = the length of the intercensal period

${}_nR_x$ = the ratio of the observed intercensal cohort survival rate to the corresponding life-table survival rate for the cohort aged x to $x+n$ years at the time of the first census

${}_nL_{x+i}$ = the life-table number of person-years lived in the age interval $x+i$ to $x+i+n$ years

${}_nL_x$ = the life-table number of person-years lived in the age interval x to $x+n$ years

and all other terms are as defined above.

In the absence of census error, the expected value of this ratio would be 1.0. Ratio values which differ from 1.0 indicate either errors in one or both censuses, distortions due to migration, or the use of an inappropriate life table. In the absence of migration and assuming the use of an appropriate life table, ratio values for any particular cohort which exceed 1.0 would indicate overenumeration of the cohort in the second census relative to the first, while ratio values of less than 1.0 would support the opposite interpretation.

6.2 *Data required*

The following information is required for the calculation of intercensal cohort survival rates:

(1) The populations enumerated in two successive censuses by age and sex.

(2) Life tables (by sex) assumed to be representative of mortality conditions during the intercensal period.

(3) Information on the volume of net migration during the intercensal period by age and sex (where applicable).

6.3 *Computational procedure*

For the procedure involving the comparison of the size of birth cohorts enumerated in successive censuses, the computational steps described below are required.

6.31 Step 1: Adjustment for migration.--In countries experiencing significant levels of net intercensal migration, the initial step in the computational procedure entails the adjustment of one of the censuses in order to minimize the distorting effects of migration on the implied cohort survival rates. A procedure for accomplishing this would be to add or subtract the estimated number of net intercensal migrants to one of the census counts. In cohorts experiencing net immigration during the intercensal period, the number of net intercensal immigrants in each cohort may either be added to the cohort population enumerated in the first census or subtracted from the cohort population enumerated in the second census. Similarly, in cohorts experiencing net intercensal emigration, the number of net intercensal emigrants can either be added to the second census population or subtracted from the first census population. More refined procedures, such as applying life-table survival rates to the estimated number of net immigrants to derive an estimate of the number of immigrants surviving to the date of the second census, also may be applied (see section 6.4 for a more detailed discussion). Adjustments for migration are illustrated below.

6.32 Step 2: Calculation of census cohort survival rates.--Having introduced adjustments for migration as necessary, cohort survival rates between the two censuses are then calculated for each sex separately or both sexes combined as in equation (5.23).

6.33 Step 3: Calculation of life-table survival rates.--Life-table survival rates reflecting the expected effects of intercensal mortality on the size of birth cohorts in the intercensal period are calculated as:

(5.25) $$ {}_nS_x = {}_nL_{x+i} / {}_nL_x $$

Where:

${}_nS_x$ = the life table survival rate for the cohort aged x to $x+n$ years

${}_nL_x$ = the life table number of person-years lived in the x to $x+n$ age interval

${}_nL_{x+i}$ = the life table number of person-years lived in the age interval $x+i$ to $x+i+n$

6.34 Step 4: Calculation of cohort survival ratios.--The final step is to calculate the ratio of the intercensal cohort survival (adjusted for migration where necessary) and the life-table survival rate for each cohort;

(5.26) $$ {}_nR_x = \frac{{}_nP_{x+i} \div {}_nP_x}{{}_nL_{x+i} \div {}_nL_x} $$

Where:

${}_nR_x$ = the ratio of the census cohort survival rate to the life-table survival rate for the cohort aged x to $x+n$ in the first census

and all other terms are defined above.

6.4 Examples

To illustrate the use of the cohort survival rates method, the method was applied to the 1960 to 1980 censuses of the Philippines and the 1881 to 1961 censuses of India.

6.41 1960 to 1980 censuses of the Philippines.--In the first application, intercensal cohort survival rates are used to assess the consistency of the size of female birth cohorts enumerated in successive censuses in the Philippines from 1960 to 1980. Three sets of cohort survival rates were calculated corresponding to the intercensal periods 1960-1970, 1970-1975, and 1975-1980, the implied cohort survival rates, calculated as shown in equation (5.23) above, are displayed in figure 5-42.

The ratios shown in figure 5-42 are the ratios of the cohort population enumerated in the second of each of the pairs of censuses considered to that enumerated in the first census of each pair, displayed for the age of the cohort in the first (reference) census. The ratio shown for the cohort aged 0 to 4 in the 1960 census, for example, was calculated as the ratio of the enumerated population aged 10 to 14 in the 1970 census to that aged 0 to 4 in the 1960 census. The other ratios shown in the table were derived in a similar fashion.

Figure 5-42. COHORT SURVIVAL RATES FOR SELECTED COHORTS IN THE PHILIPPINES FOR THE 1960-70, 1970-75, AND 1975-80 INTERCENSAL PERIODS

Sex and age in year of reference census	Year of reference census		
	1960	1970	1975
Males			
0 to 4 years	1.08	1.10	1.02
5 to 9 years	0.88	0.97	0.93
10 to 14 years	0.87	0.96	0.89
15 to 19 years	0.86	0.96	0.90
20 to 24 years	0.84	0.98	1.01
25 to 29 years	0.99	0.98	1.02
30 to 34 years	0.96	1.10	1.05
35 to 39 years	0.89	0.93	0.94
40 to 44 years	0.92	1.03	0.94
45 to 49 years	0.77	0.93	0.91
50 to 54 years	0.85	0.94	0.91
55 to 59 years	0.76	1.01	0.94
Females			
0 to 4 years	1.12	1.07	1.02
5 to 9 years	0.99	0.96	0.95
10 to 14 years	0.97	1.01	0.97
15 to 19 years	0.89	0.93	0.95
20 to 24 years	0.84	0.92	1.00
25 to 29 years	0.96	0.92	0.99
30 to 34 years	0.95	1.03	1.02
35 to 39 years	0.90	0.89	0.94
40 to 44 years	0.93	0.97	0.98
45 to 49 years	0.80	0.87	0.97
50 to 54 years	0.88	0.87	1.00
55 to 59 years	0.84	0.95	1.05

Source: 1960 census, Philippines (1963), table 5; 1970 census, Philippines (1974), Vol. II, tables 1-7; 1975 census, Philippines (1978), Vol. II, table 6; 1980 census, Philippines (1983), Vol. II, table 3.

Corresponding life-table cohort survival rates were then derived for the purpose of comparison with the cohort survival rates implied by the censuses. The procedure used in deriving appropriate life-table survival rates consisted of several steps. First, life tables were estimated for each of the census years (1960, 1970, 1975, and 1980) based upon death registration data (adjusted for under-registration), the application of selected techniques for indirect estimation, and the projection of the assumed expectation of life at birth (e_0) from 1975 to 1980 (see U.S. Bureau of the Census, 1984c, for complete documentation of the procedures and assumptions used in deriving these life tables).

Following the derivation of life tables for each of the census years, life tables corresponding to each of the three intercensal periods were derived by taking the average of the life table values from the life tables for the census years defining the end points of the respective intercensal periods. The reference point for these "intercensal life tables" was assumed to be the mid-point of the respective intercensal periods. The values of the ${}_nL_x$ life table function of the intercensal life tables derived in this manner for females are shown in the middle panel of figure 5-43. The corresponding life table survival rates for each of the intercensal periods, calculated as indicated in equation (5.25), are shown in the bottom panel of figure 5-43.

The female cohort survival rates implied by the successive censuses (from figure 5-42) and those expected on the basis of the intercensal life tables (from figure 5-43) are compared graphically in figure 5-44. While varying somewhat across intercensal periods, a general pattern of deviation of the observed from the expected cohort survival rates is observable for each pair of censuses. For example, the cohort survival rates implied for the cohort aged 0 to 4 in each of the three intercensal periods exceed those expected on the basis of the intercensal life tables, suggesting a relative underenumeration of the 0 to 4 cohort in each of the reference censuses. Similarly, the implied cohort survival rates in the age range 15 to 25 years are consistently below those expected in each intercensal period. A tendency for census cohort survival rates for the elderly population (above age 55) to exceed

Figure 5-43. LIFE-TABLE $_nL_x$ VALUES AND SURVIVAL PROBABILITIES FOR FEMALES FOR INTERCENSAL PERIODS IN THE PHILIPPINES, 1960-1980

Age	Life-table $_nL_x$ values			Life-table survival probabilities		
	1960-1970	1970-1975	1975-1980	1960-1970	1970-1975	1975-1980
0 to 4 years.......	452,608	461,182	466,868	0.96	0.97	0.98
5 to 9 years.......	437,957	449,144	456,859	0.99	0.99	0.99
10 to 14 years.....	434,173	445,696	453,773	0.98	0.99	0.99
15 to 19 years.....	431,252	443,020	451,298	0.98	0.99	0.99
20 to 24 years.....	427,312	439,329	447,835	0.97	0.99	0.99
25 to 29 years.....	421,838	434,343	443,246	0.96	0.98	0.99
30 to 34 years.....	414,584	427,961	437,469	0.95	0.98	0.98
35 to 39 years.....	405,448	419,894	430,077	0.94	0.98	0.98
40 to 44 years.....	394,466	910,065	420,959	0.94	0.97	0.97
45 to 49 years.....	381,467	398,388	410,042	0.90	0.96	0.97
50 to 54 years.....	372,944	383,952	396,350	0.85	0.95	0.95
55 to 59 years.....	344,763	364,895	373,228	0.81	0.92	0.94
60 to 64 years.....	316,934	336,637	349,864	-	-	-
65 to 69 years.....	280,066	-	-	-	-	-

Source: U.S. Bureau of the Census (1984c), tables 10 and 14.

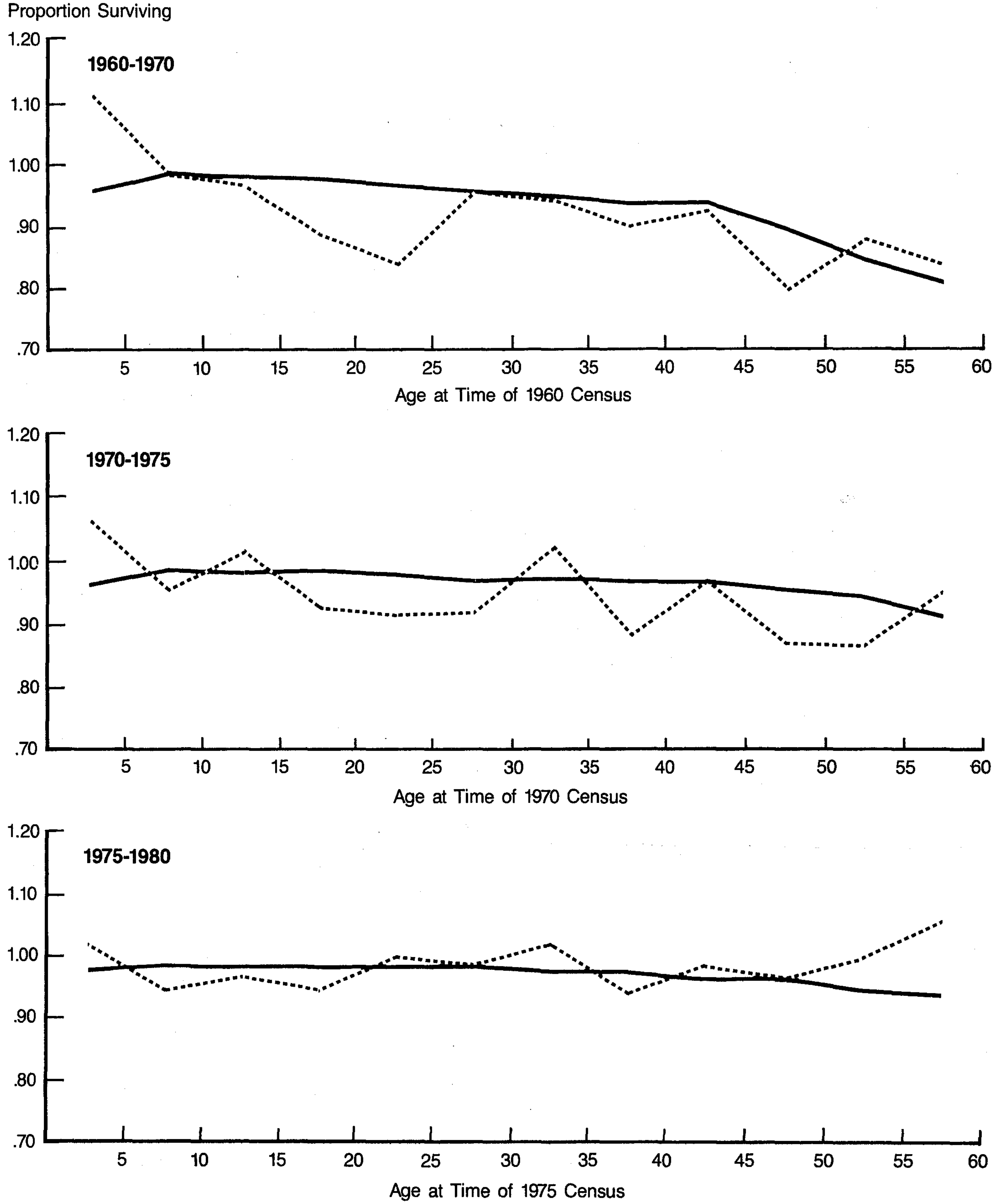
the life table survival rates is also evident in all three intercensal periods, likely reflecting exaggeration of ages among persons in this age range.

In the case of the Philippines, it is possible that the female cohort survival rates implied by successive censuses have been distorted by migration. Beginning in the early 1960's, an increasingly significant stream of female migration from the Philippines (primarily to the United States and Canada) has emerged. It has been estimated that the level of net female migration from the Philippines averaged slightly in excess of 1 percent of the female population during each year of the 1970's (U.S. Bureau of the Census 1984c).

To illustrate the use of a procedure for adjusting for the effects of migration, estimates of net cohort emigration for females during the 1975-1980 intercensal period were used to adjust the counts of female cohorts in the 1975 census. The data used in making the adjustment were derived from annual immigration statistics from the United States and Canada. Estimates of female net migration from the Philippines for the years 1975 to 1980 (by age at the time of immigration) are shown in column (1) of figure 5-45.

Prior to the use of these data to adjust the 1975 census counts, two preliminary adjustments were made. First, the estimated number of net migrants for the years 1975 and 1980 were allocated with regard to the reference date of the census (May 1 in both years). Assuming that migration was evenly distributed across months, two-thirds of the estimated number of net 1975 emigrants in each cohort (representing the months May through December 1975) were allocated to the intercensal period (that is, were assumed to have emigrated after the 1975 census date and therefore should have been enumerated in the 1975 census). Similarly, one-third of the estimated number of net emigrants in 1980 (representing the months January through April) were assumed to have emigrated during the intercensal period (that is, prior to May 1) and therefore should not have been enumerated in the 1980 census. In effect, this procedure provides an estimate of net emigration between the respective census dates.

Figure 5-44. OBSERVED FEMALE INTERCENSAL COHORT SURVIVAL RATES AND ESTIMATED LIFE-TABLE SURVIVAL RATES FOR THE 1960-70, 1970-75, AND 1975-80 INTERCENSAL PERIODS IN THE PHILIPPINES



Source: Derived from figures 5-42 and 5-43.

The resulting estimates of net intercensal emigration are shown in figure 5-45, column (2).

The second step in adjusting the migration data consisted of the allocation of the estimated number of net intercensal emigrants by age to 1975 census cohorts. These estimates shown in column (2) pertain to the age of the migrants at their time of arrival in the country of destination (and presumably at the time of emigration). However, prior to using these data to adjust the 1975 counts, the ages of the emigrants at the time of their presumed enumeration in the 1975 census must be ascertained.

Assuming that emigrants within each 5-year age group were evenly distributed with respect to single years of age, on average one-half of the net intercensal emigrants in each 5-year age group would have been enumerated in the immediately younger age group in the 1975 census. Accordingly, estimates of the number of net intercensal emigrants from each 1975 census cohort were derived by taking one-half of the net intercensal emigrants whose age at the time of emigration was the same as at the time of the 1975 census, plus one-half of the net emigrants from the adjacent older age group. For example the estimated number of net intercensal emigrants from the 1975 census cohort aged 25 to 29 shown in column (3) of figure 5-45 was derived by taking one-half of the intercensal emigrants aged 25 to 29 (.5 x 31,058 = 15,529) plus one-half of the emigrants aged 30 to 34 (.5 x 18,722 = 9,361).

These estimates of net emigration for 1975 census cohorts were then subtracted from the respective cohort populations enumerated in the 1975 census and the adjusted cohorts used in conjunction with the 1980 census counts to calculate adjusted cohort survival rates over the interval between the two censuses. The results are displayed graphically in figure 5-46. For comparative purposes, the unadjusted census cohort and life-table survival rates derived earlier are also shown.

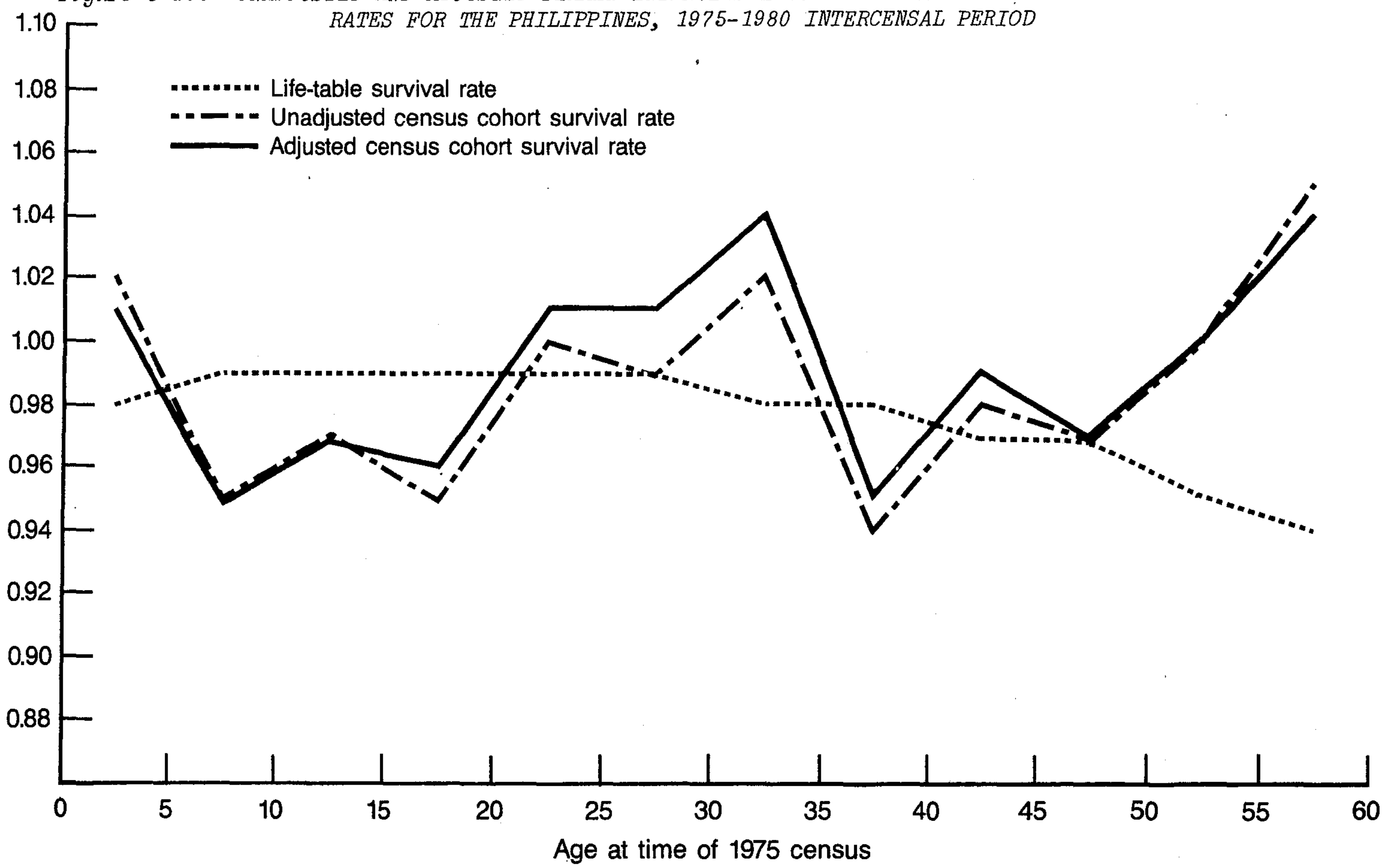
Overall, the adjustment for migration improves the "fit" of the observed to the expected survival rates only slightly, resulting in a better fit for some cohorts and a

Figure 5-45. ESTIMATED NUMBER OF NET FEMALE EMIGRANTS FROM THE PHILIPPINES DURING THE 1975-1980 INTERCENSAL PERIOD, BY COHORT

Age	Number of net emigrants (1)	Adjusted to census dates (2)	Number allocated to 1975 census cohort (3)
0 to 4 years	4,835	3,166	5,603
5 to 9 years	9,677	8,040	7,560
10 to 14 years	7,980	7,079	8,172
15 to 19 years	11,410	9,264	14,482
20 to 24 years	24,145	19,699	25,379
25 to 29 years	37,796	31,058	24,890
30 to 34 years	22,816	18,722	13,874
35 to 39 years	10,988	9,025	7,391
40 to 44 years	7,017	5,757	4,942
45 to 49 years	5,045	4,127	4,174
50 to 54 years	5,191	4,220	5,334
55 to 59 years	7,871	6,447	6,619
60 to 64 years	8,175	6,790	

Source: U.S. Bureau of the Census (1984), table 36.

Figure 5-46. UNADJUSTED AND ADJUSTED FEMALE INTERCENSAL COHORT SURVIVAL AND LIFE-TABLE SURVIVAL RATES FOR THE PHILIPPINES, 1975-1980 INTERCENSAL PERIOD



Source: Derived from figures 5-42 and 5-45.

worse fit for others. The general pattern of distortions observed in the unadjusted census data, however, remain largely intact after adjustment for the effects of migration, thereby implicating errors in one or both censuses as the major cause of these discrepancies.

6.42 <u>1881 to 1961 censuses of India</u>.--The utility of the ratio of the census cohort survival rate to the corresponding life-table survival rate defined in equation (5.26) for census evaluation purposes is illustrated in the second example using data form successive decennial censuses of India from 1881 to 1961. Cohort survival rates derived from these censuses and corresponding intercensal life tables (see Mukherjee 1976 and Ewbank 1981) were used in deriving the ratios displayed in figure 5-47. The ratio for the 0 to 4 cohort in 1881, for example, was derived by taking the ratio of the enumerated population aged 10 to 14 in the 1891 census to the enumerated population aged 0 to 4 in 1881 and dividing this result by the corresponding life-table survival rate ${}_5L_{10}/{}_5L_0$.

In the case of India, adjustment for migration was not required due to the (proportionally) insignificant levels of net migration which characterized India during this period.

The remarkable feature of the results shown in figure 5-47 is the consistency of the cohort survival rates implied by successive censuses over a period of 80 years. The observation of so consistent a pattern of discrepancies over so long a period of time constitutes strong evidence that the distortions are due to persistent errors in the successive censuses rather than cohort- or period-specific anomalies.

Figure 5-47. CENSUS SURVIVAL RATIOS FOR SELECTED COHORTS IN INDIAN CENSUSES FROM 1881-1961[2]

Sex and age in year of reference census	Year of reference census							
	1881	1891	1901	1911	1921	1931	1941	1951
Males								
0 to 4 years............	1.01	1.01	1.09	1.03	1.07	1.07	1.13	1.08
5 to 9 years............	0.72	0.68	0.71	0.65	0.69	0.69	0.76	0.81
10 to 14 years..........	1.14	1.14	1.04	1.01	1.05	1.02	1.01	1.35
15 to 19 years..........	1.58	1.42	1.44	1.33	1.50	1.36	1.35	1.26
Females								
0 to 4 years............	1.32	1.27	1.36	1.29	1.33	1.31	1.25	1.23
5 to 9 years............	0.72	0.70	0.73	0.69	0.71	0.70	0.77	0.84
10 to 14 years..........	0.83	0.81	0.79	0.76	0.80	0.82	0.83	0.85
15 to 19 years..........	1.41	1.31	1.32	1.23	1.37	1.28	1.33	1.21

Source: Mukherjee (1976:63). See also Ewbank (1981:37).
[1]*See text for definition of and computational formula for ratios shown.*
[2]*Includes only areas included in all nine censuses.*

The errors indicated by the female census life-table survival ratios for successive Indian censuses follow a pattern typical of South Asian countries (United Nations 1967; Ewbank 1981): a surplus of population aged 5 to 9, apparently caused by net age transfers from the 0 to 4 and 10 to 14 age groups; a deficit in the 10 to 14 and 15 to 19 age categories; and a large surplus in the peak childbearing ages (ages 25 to 34). The errors apparent in the enumerations of males follow a similar pattern, except for the 10 to 14 age category where a surplus rather than a deficit is observed.

The following pattern of census errors is suggested by these results: a consistent underenumeration of children aged 0 to 4, particularly of female children (as indicated by the ratios for the 0 to 4 cohort which consistently exceed 1.0 in each pair of censuses); overenumeration of females aged 5 to 9 and 10 to 14 relative to those aged 15 to 19 and 20 to 24 years; overenumeration of males aged 5 to 9 and underenumeration of those aged 10 to 14; and significant relative underenumeration of the population of both sexes aged 15 to 19.

6.5 *Uses and Limitations*

The analysis of census cohort survival rates is a widely applicable approach for examining patterns of error in successive censuses. To a large degree, the wide applicability of the method results from the fact that relatively little information other than the two census counts is needed to apply the method. Given age-sex distributions from two successive censuses, the only other piece of information required in using the method is an estimate of the level of mortality in the population which permits the selection of a model life table (in cases where an actual life table is not available). Knowledge of the level of fertility is not required since the method does not assess the coverage of the population born between the two censuses (unless three or more censuses are available). In countries which have experienced significant levels of intercensal migration, however, an estimate of the volume and age pattern of net migration during the intercensal period will prove useful in minimizing the distorting influences of migration.

When only two censuses are available, the method suffers from the limitations shared by most of the demographic methods of census evaluation, namely difficulties involved in separating census errors from other "factual" distortions on the one hand and coverage from content errors on the other. The utility of the census survival approach increases significantly, however, when three or more censuses are available. The ability to study changes in the size of birth cohorts as enumerated in successive censuses and in the relative size of cohorts born during successive intercensal periods has provided a rather firm basis for assessing the extent to which historical artifacts and demographic shifts rather than census errors have resulted in distorted distributions in several previous studies (for example, Demeny and Shorter's 1968, analyses of the 1935 to 1960 censuses of Turkey). In addition, the consistency (or lack thereof) of age patterns of apparent census errors in consecutive censuses can sometimes provide valuable clues as to the relative importance of coverage and content errors, as in the case of India described above.

The basic ideas of the census survival approach described above form the basis of a related technique for assessing the plausibility of and deriving correction factors for age distributions in two (or ideally more) successive censuses (Demeny and Shorter 1968; Das Gupta 1975; Ntozi 1978). The basic method developed by Demeny-Shorter and subsequent modifications thereof are relevant to the evaluation of content error (age misreporting), but not coverage since one of the basic assumptions of the method(s) is that the level of coverage remains approximately constant from census to census.

7. *TWO-CENSUS REGRESSION METHOD*

Preston and Hill (1980) have proposed a method which, while developed initially to provide an estimate of the degree of underregistration of deaths in the interval between two censuses, may also be used to assess the relative level of coverage in the two censuses. In many respects the method represents an extension of the basic ideas of the census survival methods described in the previous section of this chapter. In effect, the method seeks to determine coverage error correction factors for the enumerated populations (by age) in two successive censuses which, when combined with information on the number of deaths during the intercensal period derived either from a vital registration system or on the basis of a life table, results in the population of each cohort in the second census being consistent with the size of the cohort in the first census and the implied level of intercensal mortality (assuming a negligible level of net international migration). The ratio of the implied census coverage correction factors in the two censuses is used as an estimate of the relative coverage in the censuses. Ordinary least squares *(OLS)* regression procedures are used to derive the estimates of relative coverage.

7.1 *Basis of method*

The following notation, adapted from Luther (1983), is required to describe the method:

P_0 = population aged x or older as enumerated in the first census

P_1 = population age $x + i$ or older as enumerated in the second census

k_0 = correction factor for P_0 such that $k_0 P_0$ is the true population aged x or older at the time of the first census

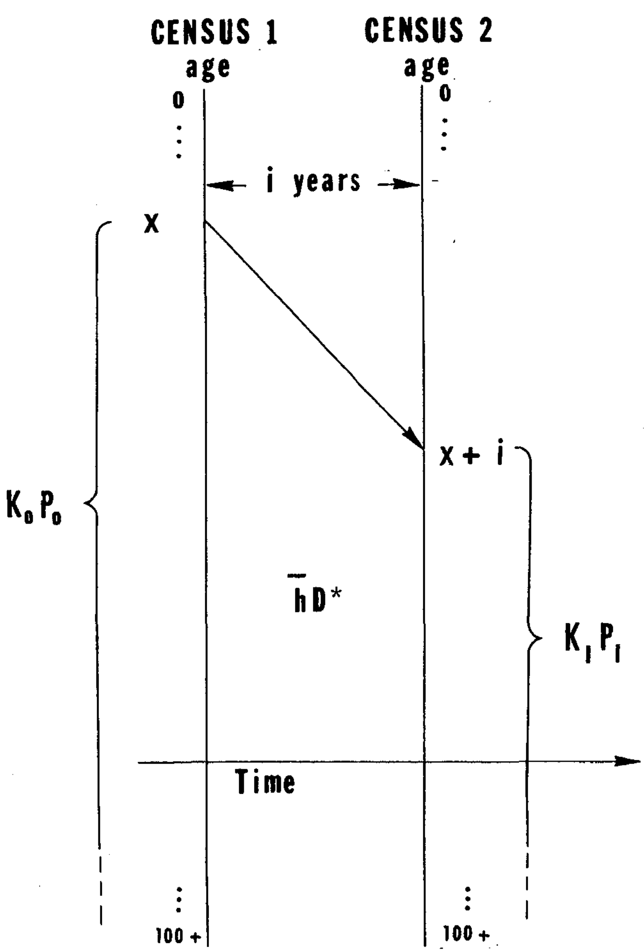
k_1 = correction factor for P_1 (such that k_1P_1 is the true population aged $x + i$ or older at the time of the second census

D^* = number of deaths registered in intercensal period to those who were age x or older at the time of the first census

$\bar{h}$ = correction factor for D^* (such that $\bar{h}D^*$ is the true number of intercensal deaths to those aged x or older at the time of the first census)

These concepts are illustrated in the lexis diagram in figure 5-48.

Figure 5-48. LEXIS DIAGRAM ILLUSTRATING THE MAJOR CONCEPTS OF THE TWO-CENSUS REGRESSION METHOD



The rationale of the method is as follows. Given two censuses taken i years apart in a country with negligible levels of net international migration, the population aged x or older at the time of the first census will be equal to the population aged $x + i$ years or older at the time of the second census plus the number of intercensal deaths in the original population (that is, of persons aged x and above at the time of the first census) or,

(5.27) $$P_0 = P_1 + D^*$$

This equality is, of course, rarely obtained in actual practice due to coverage errors in the two censuses and under-registration of deaths. If correction factors for undercoverage in the two censuses (k_0 and k_1, respectively) and for under-registration of deaths ($\bar{h}$) were obtainable, however, the equality of the expression in equation (5.27) could be obtained as follows:

(5.28) $$k_0P_0 = k_1P_1 + \bar{h}D^*$$

Equation (5.28) may be rewritten as:

(5.29) $$P_1/P_0 = -(\bar{h}/k_1)(D^*/P_0) + k_0/k_1$$

Assuming independence between the ratios $\bar{h}/k_1$, k_0/k_1 and age (see Luther 1983 for justification of this assumption), equation (5.29) may be rewritten as:

(5.30) $$P_1/P_0 = -h(D^*/P_0) + k$$

Where:

$h = \bar{h}/k_1$

$k = k_0/k_1$, and

both h and k are independent of x

Finally, if k_0 is the correction factor (which to this point remains unknown) required such that k_0P_0 equals the "true" total population at the time of the first census, the level of coverage realized in the first census may be written as P_0/P_0k_0, or $1/k_0$.

The analogous expression for the level of coverage in the second census is P_1/P_1k_1 or $1/k_1$. Redefining k in equation (5.30) as $k = k_0/k_1 = 1/k_1/1/k_0$ results in the final estimation formula shown in equation (5.31).

$$(5.31) \qquad P_1/P_0 = -h(D^*/P_0) + k$$

Where:

$$h = \bar{h}/k_1, \text{ and}$$

$$k = \frac{1/k_1}{1/k_0}$$

For census evaluation purposes, the k term in equation (5.31), which is a measure of relative level of coverage in the two censuses, is of primary interest. In the two-census regression method, k is estimated by means of ordinary least squares regression procedures. The dependent variable in the regression equation is the observed value of the ratio (P_1/P_0), the census survival rate of the population ages x and above at the time of the first census, for different values of age (x). The independent variable is the value of the ratio (D^*/P_0), the proportion of the population aged x and above at the time of the first census who had died prior to the second census, also for different values of x. A straight line is fit to the points $(D^*/P_0, P_1/P_0)$ to yield estimates of the intercept k and slope $-h$ regression parameters.

The k parameter of equation (5.31), the intercept of the regression line, is interpretable as a measure of the level of census coverage in the second census relative to that in the first census. A value for k of 1.0 would be indicative of equivalent completeness of coverage in the two censuses. Values above 1.0 would indicate improved coverage in the second census in comparison with the first, while values of less than 1.0 would support the opposite conclusion. The proportional change in level of coverage between two censuses would be calculated as $100(k-1)$.

It should be noted that although the estimate of k described above measures the difference in coverage levels between two censuses, estimates of the absolute level of coverage error in either census can be derived if either of the enumerated populations (P_0 or P_1) can be adjusted for undercoverage on the basis of other information or procedures. Thus, if the adjusted count for one of the censuses can be assumed to be accurate, the method described above will yield a direct estimate of net coverage error in the other census

The $-h$ parameter in equation (5.31), the slope of the regression line, is interpretable as a measure of the completeness of death registration during the intercensal period. This parameter will be meaningful in applications of the method based upon registered numbers of intercensal deaths, but will be less relevant when life tables are used to estimate the number of intercensal deaths.

7.2 *Data required*

The following information is required for the application of the method:

(1) counts of population by age (5-year age groups are sufficient) from two successive censuses

(2) counts of deaths by age during the intercensal period, or an appropriate life table which provides a basis for deriving age-specific survival probabilities over a period of time equal to the length of the intercensal period

7.3 *Computational procedure*

For the procedure involving the two-census regression method, the computational steps described below are required.

7.31 Step 1: Calculation of the population aged x years and above for the two censuses.--The first step in applying the method entails the calculation of a cumulative distribution of the enumerated population in each of the two censuses by age (that is, the number of persons aged x and above). Beginning with the oldest age category (75 and over, for example) in the second census, the population in each successively younger age category is added to the total population of the previous category until the initial age category (ages 0 to 4) has been included. The same procedure is then repeated for the first census, with the exception that the oldest age category chosen must reflect the length of the intercensal period. For example, if the age category of 75 and older is used as the open-ended category for the second census and the two censuses were taken 10 years apart, the appropriate open-ended category for the first census would be the population aged 65 and over. The values obtained through the procedure are the values of P_1 and P_0 (for each age x) to be used in equation (5.31).

7.32 Step 2: Calculation of the number of intercensal deaths.--The number of intercensal deaths, D^* in equation (5.31), can be obtained either from death registration data or estimated on the basis of a life table assumed to represent mortality conditions during the intercensal period. In view of its wider applicability in developing country settings, the life-table approach is described here. The reader is referred to the paper by Luther (1983) for a description of the procedures for calculating D^* based upon registered deaths.

The estimation of the number of intercensal deaths $({}_nd_x)$ to the population of each cohort enumerated in the first census is accomplished by applying the compliment of the life-table survival rate, or the "non-survival" rate, to the first census count for the cohort $({}_nP^0_x)$. Symbolically,

$$(5.32) \qquad {}_nd_x = {}_nP_x^0\left(1 - \frac{{}_nL_{x+i}}{{}_nL_x}\right)$$

Where:

i = the length of the intercensal period estimated

${}_nd_x$ = the estimated number of intercensal deaths to the cohort aged x to $x+n$ years at the time of the first census

${}_nP^0_x$ = the population enumerated in the age category x to $x+n$ years in the first census

${}_nL_x$ = the life table number of person years lived in the age interval x to $x+n$ years

${}_nL_{x+i}$ = the life table number of person years lived in the age interval $x+i$ to $x+i+n$ years

The expression $\left(1 - \frac{{}_nL_{x+i}}{{}_nL_x}\right)$ is the compliment of the life-table probability of survival, or the probability of dying, during the intercensal period.

For the oldest (open-ended) age category, the following formula is used:

$$(5.33) \qquad {}_wd_x = {}_wP^0_x\left(1 - \frac{T_{x+i}}{T_x}\right)$$

Where:

i = the length of the intercensal period

w = the highest age recorded in the census population

${}_wd_x$ = the number of intercensal deaths to persons aged x and above

${}_wP^0_x$ = the population enumerated at ages x and above in the first census

T_x = the life-table number of person years lived at age x years and above

T_{x+i} = the life-table number of person years lived at age $x+i$ years and above

The values of $({}_nd_x)$ for each cohort are transformed into values of D^* (the number of deaths in the population aged x to $x+n$ years and above in the first census) by cumulating deaths across cohorts aged $x+n$ years and above (as was done for the census populations in step 1).

7.33 Step 3: Computation of cumulated proportions dead and census survival ratios.--The cumulated census populations, P_0 and P_1, and the cumulated number of deaths, D^*, calculated in the two previous steps are then used to calculate the two ratio variables to be subjected to regression analysis. The census survival ratio (the dependent variable) is calculated P_1/P_0, while the proportion dead among the population enumerated in the first census (the independent variable) is calculated as D^*/P_0. These two ratios are calculated for each value of x.

7.34 Step 4: Calculation of regression parameters.--The final step of the computational procedure is to fit a straight line to the points (one for each value of x) with coordinates $(P_1/P_0, D^*/P_0)$ calculated in step 3 using ordinary least squares regression procedures. Since these procedures are well known and documented in numerous algebra and statistics textbooks, they are not described here. As indicated above, the intercept, k, of the regression line represents an estimate of the relative level of coverage in the two censuses, while the slope of the line, $-h$, provides a measure of the completeness of death registration (when registration data are used in deriving D^*).

7.4 *Example*

The method was applied to the 1961 and 1971 censuses of India to illustrate its application. The "smoothed" age distributions from the two censuses prepared by the Office of the Registrar General (India 1963 and 1977a) were used in lieu of the actually enumerated populations in order to reduce the variance of the regression estimates (see Luther 1983 for a comparison of results obtained using the smoothed and unsmoothed data for India). Since the adjustment procedures applied to the census age distributions did not involve adjustment of the enumerated population totals these data may be used validly to assess the relative completeness in coverage in the two censuses The population enumerated by age and sex in the two Indian censuses are shown in figure 5-49.

Following the computational procedure described above, the two census populations were first transformed into the cumulative age distributions (that is, the population at each age x and above) shown in columns (1) and (2) of figure 5-50. The values shown for the 1971 census, P_1, are listed in figure 5-50 at ages corresponding to the age of this population at the time of the 1961 census. Thus, for example, the population aged 60 and above to 1971 is shown in figure 5-50 at age 50 and above.

Figure 5-49. SMOOTHED POPULATIONS GENERATED IN THE 1961 AND 1971 CENSUSES OF INDIA, BY 5-YEAR AGE GROUPS

Age groups	Census	
	1961	1971
0 to 4 years......	72,471,900	65,045,602
5 to 9 years......	57,930,600	53,817,146
10 to 14 years....	49,667,200	46,120,355
15 to 19 years....	42,975,200	40,659,209
20 to 24 years....	37,954,100	35,921,691
25 to 29 years....	34,265,700	31,387,086
30 to 34 years....	30,322,300	26,683,910
35 to 39 years....	25,720,800	22,232,808
40 to 44 years....	24,617,700	17,891,324
45 to 49 years....	18,033,400	14,257,685
50 to 54 years....	14,620,000	10,924,614
55 to 59 years....	11,421,700	7,443,759
60 to 64 years....	8,464,200	9,876,984
65 to 69 years....	5,721,500	-
70 years or older.	7,085,200	-

Source: 1961 data reported in India (1963); 1971 data reported in India (1977a)

Figure 5-50. SMOOTHED CUMULATED POPULATIONS FROM THE 1961 AND 1971 CENSUSES OF INDIA, ESTIMATES OF CUMULATIVE NUMBER OF INTERCENSAL DEATHS, AND ESTIMATED RATIOS (D^*/P_0 AND P_1/P_0)

Age (x) at 1961 census	1961 census P_0 (1)	1971 census P_1 (2)	D^* (3)	D^*/P_0 (4)	P_1/P_0 (5)
Under 1 year........	438,271,500	382,262,173	55,398,182	.12640	.87220
5 years............	365,799,600	317,216,571	48,194,330	.13175	.86719
10 years...........	307,869,000	263,339,425	44,890,780	.14581	.85556
15 years...........	258,201,800	217,279,070	42,555,329	.16482	.84151
20 years...........	215,226,600	176,619,861	40,315,290	.18732	.82062
25 years...........	177,272,500	140,698,170	37,791,684	.21319	.79368
30 years...........	143,006,800	109,311,084	34,296,548	.23983	.76438
35 years...........	112,684,500	82,627,174	31,372,535	.27841	.73326
40 years...........	86,963,700	60,394,366	28,202,743	.32431	.69448
45 years...........	65,346,000	42,503,042	23,634,756	.36169	.65043
50 years...........	47,312,600	28,245,357	19,888,245	.42036	.59699
55 years...........	32,692,600	17,320,743	16,081,972	.49192	.52981
60 years...........	21,270,900	9,876,984	12,089,048	.56834	.46434
65 years...........	12,806,700	-	8,255,730	-	-
70 years...........	7,085,200	-	5,182,661	-	-

Source: Derived from figures 5-49 and 5-51

The second step in the procedure consists of the estimation of the number of intercensal deaths, D^*, accruing to the 1961 census population, P_0. Because of the generally poor quality of death registration data for India for the 1961-1971 intercensal period, official life tables (India 1977b) were used to estimate D^*. The values of the ${}_5L_x$ and T_x life-table functions for each age, x, from these intercensal life tables are shown (separately for each sex) in columns (3) to (6) of figure 5-51. Since life tables were available only for each sex separately, it was necessary to combine these into a single "both-sex" life table. This may be accomplished either on the basis of the sex-ratio at birth *(SRB)* in the population or by weighting the T_x values proportionally to the observed sex-ratio of the population. The former procedure is used here. Assuming a sex-ratio at birth of 105, T_x values for both sexes combined were calculated as follows:

$$T_x = \frac{(T_x \text{ (male) } \times 1.05) + T_x \text{ (female)}}{2.05} \qquad (5.34)$$

The results of these calculations are shown in column (2) of figure 5-51. The ${}_5L_x$ values shown in column (1) were derived from the T_x values on the basis of the following formula:

$$ {}_5L_x = T_x - T_{x+5} \qquad (5.35)$$

The ${}_5L_x$ and T_x life-table values shown in columns (1) and (2) were then used in conjunction with equations (5.32) and (5.33) to calculate the number of intercensal deaths to each cohort $({}_5d_x)$, which when cumulated above each age (x) as was done for the census populations $(P_0$ and $P_1)$, resulted in the D^* values shown in column (3) of figure 5-50. For reference purposes, the "non-survival" factors **applied** in equations (5.32) and (5.33) are displayed in column (7) of figure 5-51.

Following the computation of D^* values for each age, the ratios D^*/P_0 and P_1/P_0 were calculated for each age using the figures in columns (1), (2), and (3) of figure 5-50. The results of these calculations are shown in columns (4) and (5) of figure 5-50 and plotted in figure 5-52.

Figure 5-51. LIFE TABLE $_5L_x$ AND T_x VALUES AND 10-YEAR NONSURVIVAL PROBABILITIES FOR INDIA FOR THE 1961-1971 INTERCENSAL PERIOD

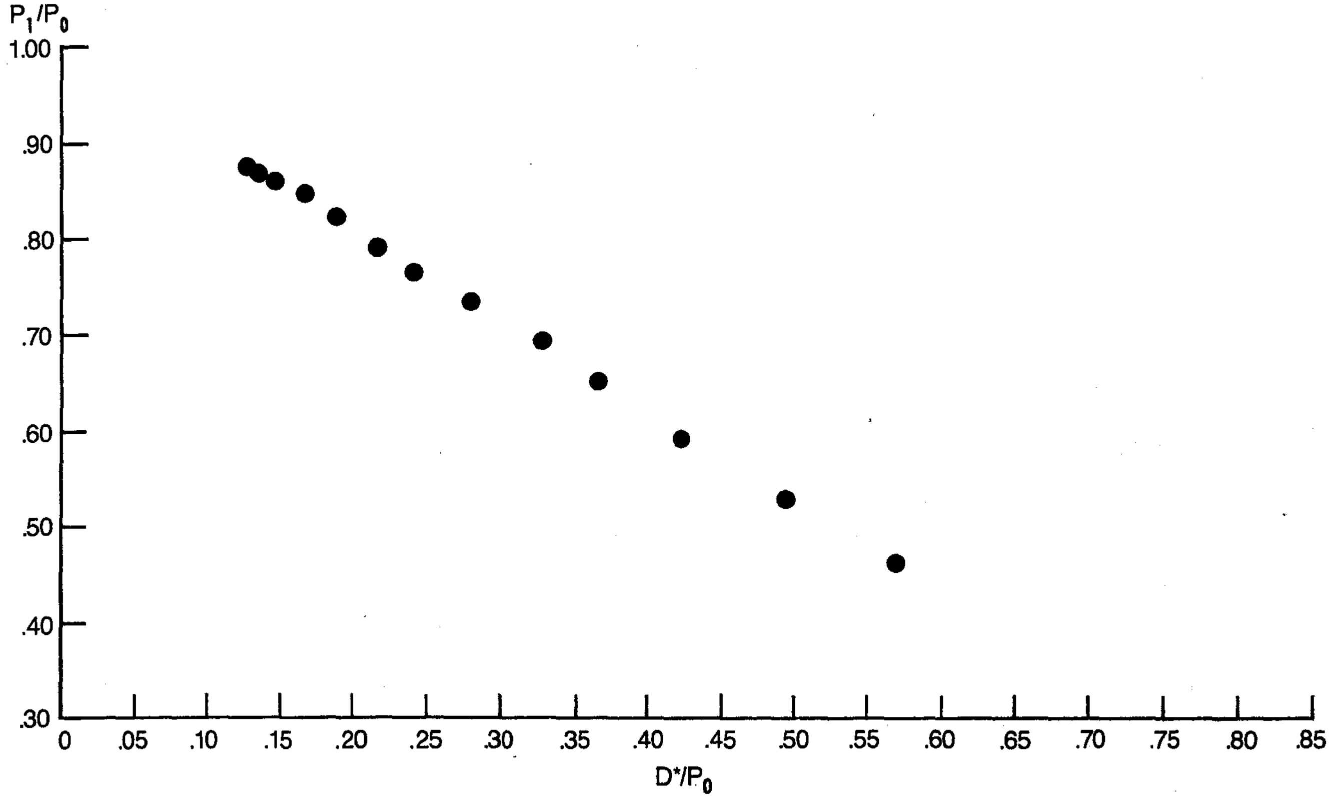
Age groups	Both sexes[a]		Males		Females		10-year nonsurvival probabilities[b]
	$_5L_x$ (1)	T_x (2)	$_5L_x$ (3)	T_x (4)	$_5L_x$ (5)	T_x (6)	(7)
0 to 4 years.......	423,369	4,553,413	423,141	4,637,190	423,609	4,465,447	.099402
5 to 9 years.......	394,919	4,130,044	397,807	4,214,049	391,885	4,041,838	.057026
10 to 14 years.....	381,285	3,735,125	386,094	3,816,242	376,235	3,649,953	.047022
15 to 19 years.....	372,398	3,353,840	376,466	3,430,148	368,127	3,273,717	.052124
20 to 24 years.....	363,356	2,981,442	366,358	3,053,682	360,203	2,905,590	.066491
25 to 29 years.....	352,987	2,618,086	355,623	2,687,324	350,220	2,545,387	.102001
30 to 34 years.....	339,196	2,265,099	343,460	2,331,701	334,718	2,195,167	.096431
35 to 39 years.....	316,982	1,925,903	318,900	1,988,241	314,970	1,860,449	.125433
40 to 44 years.....	306,487	1,608,921	321,004	1,669,341	291,244	1,545,479	.185557
45 to 49 years.....	277,222	1,302,434	288,775	1,348,337	265,090	1,254,235	.207754
50 to 54 years.....	249,616	1,025,212	260,519	1,059,562	238,168	989,145	.260347
55 to 59 years.....	219,628	775,596	228,120	799,043	210,713	750,977	.349591
60 to 64 years.....	184,629	555,968	190,606	570,923	178,351	540,264	.452886
65 to 69 years.....	142,848	371,339	147,142	380,317	138,341	361,913	.537109
70 to 74 years.....	101,013	228,491	103,926	233,175	97,953	223,572	.609911
75 to 79 years.....	66,123	127,478	67,266	129,249	64,923	125,619	.827805
80 to 84 years.....	39,404	61,355	39,878	61,983	38,906	60,696	-
85 or older........	21,951	21,951	22,105	22,105	21,790	21,790	-

Source: India (1977b), Appendix A.

[a] *T_x values for both sexes derived as follows:* $[(1.05 \times T_x^{Male}) + T_x^{Female}]/2.05$. *Values of $_5L_x$ for both sexes were derived from the T_x values.*

[b] *Derived from using the formula* $_nS_x = 1 - (_5L_{x+10}/_5L_x)$. *Probability for the 75-79 group calculated as* $_nS_x = 1-(T_{85}/T_{75})$.

Figure 5-52. PLOT OF POINTS (D^*/P_0, P_1/P_0) DERIVED FROM THE SMOOTHED AGE DISTRIBUTION FROM THE 1961 AND 1971 CENSUSES OF INDIA



Source: Derived from figure 5-50

The final step in the procedure is to fit a straight line to the points plotted in figure 5-52, whose values for each age were calculated in the previous step. The results obtained using the 1971 Indian censuses are shown in figure 5-53. The various estimates of the intercept, *k*, and slope, -*h*, parameters for each age *x* shown in figure 5-53 were obtained by fitting regression lines to successively larger numbers of points. The first estimate shown was derived based upon only two points ($x = 0$ and 5), while the last estimate takes into account all of the points up to and including the point corresponding to $x = 60$.

These estimates suggest that coverage in the 1971 census of India was somewhat poorer than in the 1961 census, although only slightly so. The estimates of differential undercoverage implied at varying points in the cumulative age distribution range from a high of about 2.8 percent for the points $x = 0,\ldots,15$, to a low of 0.6 percent for $x = 0,\ldots,55$. The median estimate of this series implies a decrease in completeness of coverage in the 1971 census of approximately 0.9 percent compared with the 1961 census, which is very close to the estimates of Premi (1982) and the U.S. Bureau of the Census (1978a).

As indicated above, the estimates generated by the method are estimates of degree of differential coverage in the two censuses considered. In order to estimate the completeness of coverage in the 1971 census on an absolute basis, an estimate of the level of undercoverage in the 1961 census is needed. If the U.S. Bureau of the Census (1978b) estimate of net underenumeration in the 1961 census of 2.7 percent is accepted, the implied net undercount in the 1971 Indian census would be 3.6 percent (2.7 percent plus 0.9 percent).

Figure 5-53. ESTIMATES OF THE LEVEL OF COVERAGE IN THE 1971 CENSUS OF INDIA RELATIVE TO THE 1961 CENSUS

Points used in Fitting Line to Estimate k and $\bar{h}$	k	$\bar{h}$
x = 0, 5	.99057	.93645
x = 0, 5, 10	.97960	.85117
x = 0, ..., 15	.97212	.79474
x = 0, ..., 20	.97732	.83256
x = 0, ..., 25	.98526	.88808
x = 0, ..., 30	.99248	.93638
x = 0, ..., 35	.99143	.92978
x = 0, ..., 40	.98894	.91485
x = 0, ..., 45	.99210	.93290
x = 0, ..., 50	.99314	.93844
x = 0, ..., 55	.99364	.94095
x = 0, ..., 60	.99228	.93460
Mean	.98741	
Median	.99100	

Source: Derived from figure 5-51.

7.5 *Uses and Limitations*

The two-census regression procedure described above provides a basis for assessing the relative completeness of coverage in successive censuses. Aside from the population counts by age (and sex if sex-specific estimates are desired) from the two censuses, the only other information required to apply the method is information on the level of mortality during the intercensal period. Either death registration data or a life table which may be assumed to be representative of mortality conditions during the intercensal period may be used for this purpose. Despite its apparent complexity, the method is quite straightforward computationally and may be applied using only a hand calculator.

While experience to date with the method is limited, the results obtained by Luther (1983) in applications to seven Asian countries suggest considerable utility in the method. In his paper, Luther shows that estimates obtained using the two-census regression procedure agree quite well with estimates for these countries based upon other methods.

With regard to limitations of the method, two points warrant mention. The first point concerns the effects of migration on the resulting estimates. Since intercensal migration will distort the observed census cohort survival rates, P_1/P_0, which are critical to the estimation procedure, adjustments to one of the two census counts will need to be undertaken in countries experiencing significant levels of net migration prior to the application of the method.

The second point concerns the sensitivity of the method to errors in the basic input data. Because of the relatively few applications of the method undertaken to date, the extent to which such factors as severe census age misreporting errors and grossly distorted mortality information used in deriving D^* adversely affect the resulting estimates of differential census coverage in successive censuses is somewhat uncertain. For example, Luther observes in his paper that the estimates of differential coverage in the 1961 and 1971 Indian censuses obtained using the unadjusted census counts vary considerably from age to age (that is, at different values of x). While the overall estimate of the degree of differential coverage in the two censuses obtained on the basis of the unadjusted count is close to that obtained using the smoothed age data, the variance of the estimate is significantly higher.

In short, while the method appears to have performed adequately in previous applications, further investigation of the robustness of the method is warranted.

8. REFERENCES

Argentina, Instituto Nacional de Estadística y Censo. 1982. *Censo Nacional de Población y Vivienda 1980*. Buenos Aires, Instituto Nacional de Estadistica y Censo.

Arnold, F., and Phananeramai, M. 1975. *Revised Estimates of the 1970 Population of Thailand*. Research Paper No. 1. Bangkok: National Statistics Office.

Bangladesh Bureau of Statistics. 1983. *1981 Bangladesh Population Census Post Enumeration Check.* Dhaka: Bangladesh Bureau of Statistics.

Bourgeios-Pichat, T. 1974. *An Attempt to Appraise the Accuracy of Demographic Statistics for an Under-Developed Country: Thailand.* Perspectives on Thai Population, Research Report No. 11. Bangkok: Chulalonghorn University, Institute of Population Studies.

Brass, W. 1975. *Methods for Estimating Fertility and Mortality from Limited and Defective Data.* Chapel Hill: Carolina Population Center and Laboratories for Population Studies.

Chandrasekar, C., and Deming, W.E. 1949. "On a Method of Estimating Birth and Death Rates and the Extent of Registration." *Journal of the American Statistical Association* 44(249):101-115.

Coale, A., and Demeny, P. 1966. *Regional Model Life Tables and Stable Populations.* Princeton: Princeton University Press.

Coale, A.J., and Rives, N.W. 1973. "Statistical Reconstruction of the Black Population of the United States 1880-1970: Estimates of Time Numbers by Age and Sex, Birth Rates, and Total Fertility." *Population Index* 39(1):3-36.

Coale, A.J., and Zelnick, M. 1963. *New Estimates of Fertility and Population in the United States.* Princeton: Princeton University Press.

Concepcion, M.B. 1974. *Changes in Period Fertility as Gleaned from the 1973 NDS.* Manila: University of the Philippines.

Das Gupta, P. 1975. "A General Method of Correction for Age Misreporting in Census Populations." *Demography* 12(2):303-312.

Demeny, P., and Shorter, F. 1968. *Estimating Turkish Mortality, Fertility, and Age Structure.* Istanbul: Istanbul University.

Gardiner, P. 1978. *Age Patterns of Mortality in Indonesia: An Analysis of Results of the Indonesian Sample Vital Registration Project, 1974-1977.* Honolulu: East-West Center.

India, Office of the Registrar General. 1963. *Census of India 1961.* Age Tables, Paper No. 2. New Delhi: Office of the Registrar General.

-----, Office of the Registrar General. 1977a. *Census of India 1971.* Age Tables, Series I, Paper No. 3. New Delhi: Office of the Registrar General.

-----, Office of the Registrar General. 1977b. *Census of India 1971.* Life Tables, Series I, Paper No. 1. New Delhi: Office of the Registrar General.

-----, Registrar General and Census Commissioner for India, 1982. *Census of India 1981.* Report on Post Enumeration Check, Series I, Doc. No. 4. New Delhi: Registrar General and Census Commissioner for India.

Indonesia, Central Bureau of Statistics. 1963. *Population Census 1961, Preliminary Figures Extended 1% Sample.* Series Sp. 2. Jakarta: Central Bureau of Statistics.

-----, Central Bureau of Statistics. 1975. *Population Census 1971.* Population of Indonesia, Series D. Jakarta: Central Bureau of Statistics.

-----, Central Bureau of Statistics. 1976 *Estimates of Fertility and Mortality Based on the 1971 Population Census.* Jakarta: Central Bureau of Statistics.

-----, Central Bureau of Statistics. 1981. *Fertility and Family Planning in Indonesia: Results of the National Socio-Economic Survey, 1979.* Volume I. Jakarta: Central Bureau of Statistics.

-----, Central Bureau of Statistics. 1982. *Levels of Births and Deaths Based on the 1971 Census, 1976 SUPAS, and 1980 Population Census.* Jakarta: Central Bureau of Statistics.

-----, Central Bureau of Statistics. 1983. *Population of Indonesia: Results of the 1980 Population Census.* Series S, No. 2. Jakarta: Central Bureau of Statistics.

Japan, Statistics Bureau. 1982. *Results of the First Basic Complete Tabulation, Part I. Whole Japan.* Volume 2 of *1980 Population Census of Japan.* Tokyo: Statistics Bureau.

Krotki, K., ed. 1978. *Development in Dual System Estimation of Population Size and Growth.* Alberta, Canada: University of Alberta Press.

Luther, N. 1983. "Measuring Changes in Census Coverage in Asia." *Asian and Pacific Forum* 9(3):1-16. Honolulu: East-West Center.

Luther, N.Y., and Retherford, R.D. 1985. "Consistent Correction of Census and Vital Registration Data." Paper presented at the Annual Meeting of the Population Association of America, Boston.

Marks, E.S., and Rumford, J.C. 1978. "The 1974 Post Enumeration Survey of Liberia--A New Approach." In *Developments in Dual System Estimation of Population Size and Growth,* edited by J.K. Krotki, 189-204. Edmonton: University of Alberta.

Mijares, T.A. 1974. *Development and Maintenance of the Sample Vital Registration System in the Philippines.* Volume 23 of the *Philippines Statistician.* Manila: Philippines Statistical Association.

Mukherjee, S.B. 1976. *The Age Distribution of the Indian Population.* Honolulu: East-West Center.

Ntozi, J.P.M. 1978. "The Demeny-Shorter and Three-Census Methods for Correcting Age Data." *Demography* 15(4):509-522.

Pagtolun-an, Imelda G. 1983. "Rural-Urban Mortality Differences in the Philippines." Unpublished paper.

Pakistan, Office of the Census Commissioner. n.d. "Census of Pakistan, 1931." *Census Bulletin* No. 5: Parts I and II. Karachi: Office of the Census Commissioner.

-----, Office of the Census Commissioner. n.d. *Census of Pakistan.* Vol. 1 of *Population 1961.* Karachi: Office of the Census Commissioner.

-----, Population Census Organization. 1979. Unpublished Computer Printouts of 1972 Census Data. Islamabad: Population Census Organization.

-----, Population Census Organization. 1981. "Housing and Population Census of Pakistan 1980-1981." *Census Bulletin* No. 2. Islamabad: Population Census Organization.

Philippines, National Census and Statistics Office. 1963. *1960 Census of the Philippines.* Vol. 2 of *Summary Report.* Manila: National Census and Statistics Office.

-----, National Census and Statistics Office. 1974. *1970 Census of Population and Housing.* Vol. 2 of *Summary Report.* Manila: National Census and Statistics Office.

-----, National Census and Statistics Office. 1978. *1975 Integrated Census of Population and Its Economic Activities.* Vol. 2 of *National Summary.* (Phase I). Manila: National Census and Statistics Office.

-----, National Census and Statistics Office. 1979. *Republic of the Philippines Fertility Survey.* Manila: National Census and Statistics Office.

-----, National Census and Statistics Office. 1983. *1980 Census of Population and Housing.* Vol. 2 of *National Summary.* Manila: National Census and Statistics Office.

Premi, M. 1982. "The Demographic Situation in India." Doc. No. 80 of *Papers of the East-West Population Institute.* Honolulu: East-West Center.

Preston, S.; Coale, A.J.; Trussell, T.J.; and Weinstein, M. 1980. "Estimating the Completeness of Reporting of Adult Deaths in Populations that are Approximately Stable." *Population Studies* 34(2):170-202.

Preston, S., and Hill, K. 1980. "Estimating the Completeness of Death Registration." *Population Studies* 34:349-366.

Sri Lanka, Department of Census and Statistics. 1953. *Post Enumeration Survey 1953.* Monograph No. 1. Colombo: Department of Census and Statistics.

-----, Department of Census and Statistics. 1970. *A Study of the Extent of Under-Registration of Births and Deaths in Ceylon.* Colombo: Department of Census and Statistics.

-----, Ministry of Plan Implementation, Department of Census and Statistics. 1982. *Statistical Abstract of the Democratic Socialist Republic of Sri Lanka.* Colombo: Ministry of Plan Implementation, Department of Census and Statistics.

Suharto, S., and Cho, L.S. 1978. *Preliminary Estimates of Indonesian Fertility Based on the 1976 Intercensal Population Survey.* Doc. No. 52 of *Papers of the East-West Population Institute.* Honolulu: East-West Population Institute.

Thailand, National Statistics Office. 1969. *Report: The Survey of Population Change 1964-1967.* Bangkok: National Statistics Office.

-----, National Statistics Office. 1973. *Whole Kingdom, 1970 Census of Population.* Bangkok: National Statistics Office.

Trussell, J., and Menken, J. 1979. "Estimating the Completeness of Death Registration and the Relative Underenumeration in Two Successive Censuses." *Asian and Pacific Census Forum* 6:9-11.

United Nations. 1955. *Methods of Appraisal of Quality of Basic Data for Population Estimates*. Manual 2. New York: United Nations.

-----. 1967. *Methods of Estimating Basic Demographic Measures from Incomplete Data*. Manual 4. New York: United Nations.

-----, Economic and Social Commission for Asia and the Pacific. 1978. *Population of the Philippines*. Country Monograph Series No. 5. Bangkok: United Nations.

-----, Population Division, and the U.S. Academy of Sciences. 1983. *Indirect Techniques for Demographic Estimation.* Manual 10. New York: United Nations Population Division and the U.S. National Academy of Sciences.

-----, Department of International Economic and Social Affairs. 1984. *Population and Vital Statistics Report: 1984 Special Supplement.* New York: United Nations.

U.S. Bureau of the Census. 1975. *The Methods and Materials of Demography.* Washington, D.C.: U.S. Bureau of the Census.

-----. 1975. *Levels and Trends of Mortality in Indonesia: 1961-1971.* International Research Doc. No. 2. Washington, D.C.: U.S. Bureau of the Census.

-----. 1977. *Country Demographic Profiles: Sri Lanka.* Country Demographic Profile Series, Doc. No. 8. Washington, D.C.: U.S. Bureau of the Census.

-----. 1978. *Country Demographic Profiles: Thailand.* Country Demographic Profile Series, Doc. No. 15. Washington, D.C.: U.S. Bureau of the Census.

-----. 1978. *Country Demographic Profiles: India.* Country Demographic Profile Series, Doc. No. 16. Washington, D.C.: U.S. Bureau of the Census.

-----, Center for International Research. 1984. *Detailed Statistics on the Urban and Rural Population of Indonesia: 1950 to 2000.* Washington, D.C.: U.S. Bureau of the Census.

-----, Center for International Research. 1984. *Detailed Statistics on the Urban and Rural Population of Pakiston: 1950 to 2000.* Washington, D.C.: U.S. Bureau of the Census.

-----, Center for International Research. 1984. *Detailed Statistics on the Urban and Rural Population of the Philippines: 1950 to 2000.* Washington, D.C.: U.S. Bureau of the Census.

Yemen Arab Republic, Statistics Department. 1980. *Evaluation and Analysis of the 1975 Population and Housing Census: Testing the Accuracy of Age-Sex Statistics.* Yemen Arab Republic: Statistics Department.

Chapter 6. ADJUSTING CENSUS FIGURES

1. INTRODUCTION

Earlier chapters reviewed the types of errors that can occur in a census. If the errors are substantial and the validity of the counts delivered by the census is in question, the central statistical office may wish to consider adjusting the census counts using information from the evaluation studies. This chapter will examine the questions raised in considering an adjustment: how does one decide whether to make an adjustment, what effects will the adjustment have on the data collected and other products from the census, what procedures are available for making an adjustment, and how does one choose among alternative adjustment procedures?

The implications for making an adjustment of the census figures are so far-reaching, however, that the decision to do so should be made at the time general census planning is taking place, rather than after the evaluation results have been processed. For example, the decision to make an adjustment will require the commitment of additional resources to the evaluation of the census. For a statistical office with a fixed budget, this means financial resources and staff are shifted away from the conduct of the census. In addition, the decision to adjust the data leads to choices between producing census counts in a timely manner and obtaining population estimates of the highest quality. A tradeoff arises in the choice between the timing of the census and the degree to which the counts can be improved. Decisions of this sort can only be addressed for each census individually; there is no general guideline for deciding whether to make an adjustment or for resolving the tradeoffs raised by the commitment of resources in the census.

This chapter will show that the issues surrounding census adjustment are varied and complex. Moreover, not many countries have actually adjusted their census figures, so past experience is very limited in terms of guiding the decision process. For these reasons, the chief benefit of this chapter will be in focusing attention on what the issues are, not necessarily what the best solutions might be.

2. GENERAL CONSIDERATIONS IN DECIDING TO MAKE AN ADJUSTMENT

2.1 *The initial decision to make an adjustment*

The decision whether to adjust census counts is very likely to be made at the highest levels of government; for example, at the ministerial level. Though there are many technical considerations that must be taken into account, there are also policy consequences that will undoubtedly affect the choice. Making the decision can be a sensitive process, for it is quite possible that policy concerns could override the technical ones. In any case, it will not likely be the analysts, statisticians, or technicians alone who ultimately decide whether an adjustment will be done; it is their responsibility, based on their technical expertise, to make appropriate recommendations to the decision-makers.

On the technical level, there is a dilemma not usually encountered in other evaluation studies. In other evaluation settings, it is usually possible to determine, for a subsample of the population, a more accurate measure of the variable being studied. But for the estimation of a population total, it may be impossible to get a measure of the total which can be presented with confidence as being more accurate than the census count. The analyst using estimates from a PES or demographic analysis may only be able to say that the estimates are an alternative to the census count. Deciding whether to use the census count or an estimate of the total from a research study designed for such a purpose has to be based on a theoretical justification for use of the count or an estimate. There is no way to know the true population total, and so there is no way to be sure that the adjustment using an estimate of the population gets one closer to the true population total.

Another issue is in deciding where and how to make an adjustment. If an adjustment is made, it would cover a variety of geographic detail, and have an effect on demographic distributions. Consideration will have to be given to what levels of geography and which demographic variables are most important in the decision to make an adjustment. With this information and the theory underlying the population estimation procedures, the analyst can recommend whether to make an adjustment and which procedure to use.

The analyst will need to know how to calculate gain or loss from making an adjustment and also needs to develop a rule for determining whether an adjustment should be made. These should be chosen before analyzing the data available for the adjustment. The choice of an objective function as described above should reflect the effect of the adjustment on the uses of the published census data, and the choice of a decision rule should reflect the cost of making an adjustment. Examples will help explain the choices to be made.

Two objective functions that might be considered would be an adjustment that reduced the average error in the population totals versus one that would minimize the maximum error in the estimates. The first alternative is one which in expectation will minimize the average error in the estimates, but for which some of the changes due to adjustment can be quite substantial. The second alternative procedure would keep the largest errors after adjustment to a minimum, but might allow all errors to be larger on average. Since one will never know the errors in the estimates after adjustment of the census counts, one would use a proxy variable to indicate average or maximum error. Such a proxy might be the degree of change from the census counts, using the adjusted figures, as measured by the proportional change from the census count. If there is reason to believe that some geographic areas or some demographic subgroups were especially poorly covered in the census, then the degree of change may serve as a poor proxy, because minimization of change from the census would defeat the purpose of making the adjustment.

On the other hand, if the coverage of the census is low but fairly uniform across all geographic and demographic subgroups, then one may recommend against making an adjustment. As an extreme example, suppose each area covered by the census was undercounted by 10 percent. If the primary use of the census is for the disbursement of funds and resources by the national government based on proportional allocation, an increase of 10 percent in the population totals for each geographic area would have no effect on funds disbursement. There would be no benefit to making an adjustment, though the cost of doing it could be substantial.

To summarize, in order to make appropriate recommendations to the decision-makers, the analyst is faced with several choices from the outset. The first is whether to make an adjustment: a decision rule, or set of decision rules is needed to indicate whether the adjustment will make a sufficient difference in the final estimates of the population to compensate for the cost of making the adjustment. It may be that the estimated undercount is too small or not variable enough across geographic areas to make an adjustment worthwhile. The second choice arises if more than one adjustment procedure is available. In this case the analyst must have an objective function to choose between procedures, or a function that would allow the combination of a set of population estimates.

2.2 *Specific issues in considering adjustment*

If an initial decision is made to adjust, the statistical office must determine how to do it. Decisions must be made about what counts should be adjusted, the geographic level at which adjustment is to be made, the consistency of the estimates, and the extent of the information to be adjusted. Most of these issues revolve around the publication and presentation of the counts from the census. The statistical office making the adjustment has the option of publishing both the census counts and the estimates from the adjustment, or only the adjusted counts. If both are published, it will have to be made clear at the time of publication which set is considered the official census figures. Failure to designate official census figures will lead to a variety of problems following publication, since different users would likely choose the set of numbers most advantageous to their vested interests.

A second decision is needed regarding what data are actually to be adjusted. For example, a census collects information on population, housing, and other characteristics. If one were to adjust only the population counts, but not the housing counts, figures such as the average number of persons per housing unit might become useless, and hence this index as an indicator of housing conditions would no longer be valid. A related issue is when different characteristics are to be adjusted, whether different procedures should be used to adjust the characteristics, or for consistency, whether one procedure should be used to make an adjustment for all characteristics.

Various procedures can be used to estimate the level of undercount in the census, but each procedure (e.g., a PES) is usually applicable to some minimum level of geography. Below that level, some other technique (e.g., regression) must be used to allocate the undercount, or a decision must be made not to allocate the undercount to lower levels of geography. If the latter course is chosen the statistical office is faced with the problem of publishing figures for low levels of geography, such as blocks, that will not sum to published adjusted figures at higher geographic levels. Even if an adjustment is allocated to lower levels of geography, the technique chosen will not necessarily provide estimates at all lower levels that would sum to adjusted counts, unless the procedure carefully takes into account the total publication plans of the statistical office. See section 5.2 of this chapter for more information on allocating the count to lower levels of geography.

On the matter of consistency, it may be that there is more information available for certain portions or segments of the population, so that a different allocation procedure could be used for these segments than for others. The analyst has to choose between using the additional information that is available and the consistency of using only one procedure across all segments of the population, even though the latter procedure may be optimal for each segment.

The last issue to be addressed in this section is the extent of the adjustment. The notion of adjustment usually implies a census undercount - persons completely missed who should be counted. A narrower definition of adjustment, however, would be the use of imputation - that is, to make corrections for missed persons about whom there is some evidence that a person (or at least a possible residence) was missed. For example, if there are some residences where no person has been contacted, the analyst may wish to impute persons into these residences, using other persons in the same enumeration area as the "deck" from which to impute. Often there will be information as to the number of persons and composition of the household and this information can be used as part of the imputation process. In the case where no information is available other than the fact that a residence exists but has not been enumerated, the number of persons is imputed first, and then the composition of the household is imputed from neighboring units.

The latter approach is more restrictive than an adjustment for the undercount because it requires some tangible evidence that a unit or person was missed. But additional adjustment, beyond imputation, can also be based on a PES approach, where, through a matching study, other units and persons are found to be missed from the census, but not through the exercise of the regular census procedures. Furthermore, it may be acknowledged that the frame for the survey used for the PES is incomplete, so an estimate is made of persons missing from the PES frame; this estimate is used as part of the adjustment process.

The broadest possible adjustment would also take account of an estimate of the overcount in the census. Duplicate enumerations and persons who should not have been enumerated (for example, children born after census day) could be deducted from the census total to give a more accurate estimate of the correct number of persons in the population. Because this problem is traditionally much less severe than problems of undercounting, there has been a tendency not to make an adjustment for overcounts. Duplicates and erroneous enumerations are harder to estimate correctly, can be scarce relative to the undercounted population, and so may add unwarranted difficulties to the adjustment process.

3. *IMPACT OF VARIED USES OF CENSUS DATA*

As part of the analysis of whether and how to make an adjustment of the census, the analyst has to consider what uses are made of census data. An important use for the data is program implementation and funds distribution. Some national and local government agencies rely on census figures to determine how to allocate funds and resources to various government programs. The question for the analyst is to find out what data are mandated for use, what programs are most severely affected, and what variables are most often used to determine funds allocations.

As pointed out in the previous section, adjusting only the population counts may do more harm than good if variables such as number of persons per household, per capita income, and other ratio variables are predominant in funding formulas. The adjustment process must consider each of the variables used in funds allocation, and for variables which will be adversely affected by the adjustment, and offer alternative or proxy variables for the program affected.

Another use of census data is in program management. The question of timeliness is critical here. While the adjustment may actually improve the quality of the census data used in program management, if it lags too far behind the census it would be of little use for some programs. This may indicate that the adjustment contemplated should be made less comprehensive in order to meet the timing requirements for some important programs. Besides restrictions on the timing of the production of counts, there could also be regulations against the use of statistical or adjusted data for some programs. In this case two sets of census estimates would be produced - adjusted and unadjusted.

Besides being used for the distribution of program funds, in some countries the census counts may be used to define political boundaries and to apportion seats in the national legislature. The apportionment process must often be completed soon after the census, again constraining the adjustment in terms of the time available to complete the research necessary. In some cases the apportionment of the legislature is based solely on population counts for existing geographic areas. In others, geographic areas are actually redefined based on census population counts and a variety of other census variables (called redistricting). Again, the consideration here is how variables other than the census counts are affected by adjustment and what the effect is on the redistricting or apportionment process.

In addition to funds distribution and uses for defining legislative boundaries or apportionment, there are many other uses of census data that need to be taken into account for adjustment purposes. One use of the census is as a sampling frame for surveys. Any redesign of the sampling frame would be affected by an adjustment in several ways - the definition of strata, the combination of geographic entities into primary sampling units, and the determination of probabilities of selection. Following the conduct of a survey, the adjusted census data could be used to poststratify the sample or to develop weighting factors for ratio or raking adjustments.

Besides survey uses, private businesses and researchers in other governmental agencies will have other uses for census data. For example, census data are commonly used to determine the location of new hotels, shopping centers, and government service centers. Census data are used to determine and improve traffic flows or to analyze marketing strategies for various products. The analyst, in considering these uses, needs to determine the variables used for these applications, the effects an adjustment would have, and whether any of these applications require unadjusted data.

4. OPERATIONAL CONSIDERATIONS IN MAKING AN ADJUSTMENT

Before an adjustment is made, certain operational factors must be considered. These include when to make an adjustment, the type of revisions to be issued (if any), and how the adjustment process relates to the main census.

4.1 *Timing and revisions*

One of the most important considerations in making an adjustment is that it be done on a timely basis. The best adjustment procedure is useless if it takes too long to perform relative to the demand for census data. The analyst has to determine when the counts are needed and what uses of the census will suffer if delivery of the counts is delayed by an adjustment.

Anticipated early uses of the census data may severely restrict the time available for adjustment based on complete evaluation results. So a provisional adjustment could be made, which would be revised as more evidence became available from the evaluation program, or a series of revisions could be incorporated. With this strategy, the analyst would need to know when the revisions would be needed to fit into various programs, and when revisions would be least disruptive to the publications program for the census. The decision would also have to be made as to how to publicize the revisions. Publications from the census could be reissued and if computer tapes and geographic products like maps are also issued following the census, these could also be revised. But the revision process is costly and confusing. The more estimates produced, the harder it will be to disseminate the information to all users of census data.

The decision to make a revision to the adjustment is closely akin to the decision initially to make the adjustment. The same types of questions need to be asked: whether the change introduced by the revision is substantial enough to justify the cost of making the revision, and if more than one type of revision is plausible, how does one choose between alternative revisions.

Finally, the choice to issue revisions of the adjustment will have an effect on staffing needs. It will be necessary to advertise and document changes due to revisions. Shifting existing staff to take care of these functions or adding new staff is necessary to help users adapt to the revisions and maximize their use of census data.

4.2 *Integrating the adjustment into the census process*

As part of the consideration of timing and publication, the adjustment will affect many facets of the census processing, especially if a PES is conducted. If the statistical office wishes to make the adjustment to the census counts soon after or at the time the census is published, the adjustment process must be incorporated into the conduct of the census. The publication program, the clerical processing of the data, and computer operations all need to be changed to incorporate procedures for the adjustment. Note that this is not the same as saying the procedures must be changed to accommodate the adjustment. For the adjustment process to work, it must be considered an integral part of the census, not as an add-on feature that can be put aside if problems arise in the census.

As an example, if a PES is conducted shortly after the census, a large matching operation must be undertaken using census questionnaires. But these same questionnaires are being processed for the census: they are undergoing editing, keying, or some other process to get them on computer tapes, and the other census materials are in use at the same time to ensure a complete enumeration. If the PES is going to be used to adjust the census, some routing system has to be devised so that the PES has access to census questionnaires and materials shortly after the census. The routing system must be designed so as not to create obstructions in the census processing. A delay in census processing will eventually lead to a delay in the PES, starting a cycle of slowdowns that could eventually threaten publication of a completed product, especially if the adjustment relies on PES data.

5. METHODOLOGICAL CONSIDERATIONS IN MAKING AN ADJUSTMENT

There are many techniques available to make population estimates and to estimate the undercount in the census. These include demographic analysis, reverse record check studies, administrative record studies, and post enumeration studies. These techniques have been described either in other chapters of this document or other texts on evaluation, and will not be described further here.

5.1 *Measurement techniques*

Depending on the range of the evaluation program associated with the census, the statistical office conducting the census may carry out more than one type of study to evaluate the census. If this is the case, there will be a multiplicity of estimates available upon which to base an adjustment of the population totals, and the statistical office will want to give serious consideration to combining the estimates. Combining the estimates has the advantage of taking the best characteristics from each evaluation program and using these characteristics to counterbalance weaknesses in the other evaluation programs. For example, estimates from demographic analysis may only provide national totals, but those may be considered superior to estimates from a PES. The PES, however, can provide more geographic detail than demographic analysis, if the sample size is big enough. If the analyst believes that the differences for coverage among geographic areas are well represented in the PES, but that the totals calculated from demographic analysis are of better quality, the PES totals could be ratio adjusted to the demographic analysis totals, combining the best features of both estimates.

One other consideration is the measurement of populations which are especially hard to enumerate. An adjustment may be hampered by the lack of data for certain subgroups if these groups are not given special consideration from the beginning of the research. Persons living in group quarters, tourists, nomadic populations, and other groups may cause special problems in either measuring the undercount or allocating the undercount to defined geographic areas. Careful thought should be given to identifying these groups and how they are to be measured. If this category includes the institutional population, such as persons in prison or the military, the analyst should consider whether a separate estimate is needed of the institutional population, and the possibility for use of administrative records to aid in estimation.

5.2 *Methods of adjustment*

Just as there are several techniques available for measurement of the census undercount, there are also several techniques available for making an adjustment for lower levels of geography and more specific demographic subclasses. The techniques described in the preceding sections will only supply estimates of the undercount for fairly gross levels of geography, like regions of a country. To make an adjustment of the census and to allocate the revised population figures down to lower levels of geography, techniques which will use the data from the measurement studies can be implemented. These techniques are synthetic estimation, regression, and borrowed strength techniques. Each of these require that a set of variables, primarily demographic, be available for lower levels of geography, in some instances down to the person or household level. The techniques permit modeling the distribution of the undercount at the level of geography appropriate to the measurement technique; the model obtained is then used to allocate the undercount to lower levels of geography or to areas not in the sample.

The modeling process for synthetic estimation is to estimate the mean undercount rate (persons missed as a percent of total estimated population) for various demographic subgroups at a certain geographic level. These rates would then be used at the lowest level of geography, like enumeration area clusters, to allocate the undercount. This methodology takes the undercount at high levels of geography and distributes it proportionately at lower levels of geography. Synthetic estimates created in this fashion are guaranteed to sum to estimates of undercount generated at higher levels.

As an example, suppose that undercount estimates, U_{ij}, are available at the national level for age, i, and sex, j. Synthetic estimates, $\hat{U}_{ijk}$, would be made for each block, k, by multiplying the national level estimates by the proportion of persons having particular age/sex characteristics in the block relative to all persons having the same age/sex characteristics $\frac{X_{ijk}}{\sum_k X_{ijk}}$:

$$\hat{U}_{ijk} = \frac{X_{ijk}U_{ij}}{\sum_k X_{ijk}}$$

where the X variable is the person count.

The second technique, regression, fits a regression model to the undercount estimates for a set of geographic areas (possibly primary sampling units, PSU's, or the larger geographic units they collectively represent). The estimates are generated in a fashion similar to that used for synthetic estimation, by applying the coefficients estimated at a grosser level of geography to characteristics and variables observed in finer detail. If undercount estimates, U_i, are available within regions, i, a regression can be run on a variety of variables, like proportion of population under age 40, X_{1i}, proportion of male population in region i, X_{2i}, and proportion of census data imputed, X_{3i}, with the equation

$$U_i = a + b_1X_{1i} + b_2X_{2i} + b_3X_{3i} + e_i$$

where

a = the intercept of the regression

b_1, b_2, and b_3 = slope coefficients for the age, sex, and imputation variables respectively

e_i = the error term in the regression for region i

To go to lower levels of geography, such as enumeration areas, j, within regions, the allocation would be implemented as

$$\hat{U}_{ij} = a + b_1X_{1ij} + b_2X_{2ij} + b_3X_{3ij}$$

These estimates are not guaranteed to sum to the survey or demographic estimates formed for each region; instead the estimates $\hat{U}_{ij}$ will sum to the predicted values obtained from the regression conducted at the regional level.

An alternative regression technique would be to fit the regression to sample units, such as PSU's rather than to broad regional areas for which estimates can be obtained from the combination of several PSU's. The regression equation is used then to allocate the undercount to all units in the population, based on the regression conducted on the sample units at the same level of geography. These estimates can then be allocated to lower levels of geography using the same methodology described in the preceding paragraph.

Another allocation technique which will not be fully described here is the use of borrowed strength estimators. This technique requires ascribing a multivariate density to the undercount and its correlates; the multivariate normal is most often used. The use of prior and posterior densities in this estimation is difficult to describe without also discussing the broader category of Bayesian techniques, which is beyond the scope of this manual. In the case where the multivariate normal is used as an asymptotic approximation for describing the density of the undercount and its correlates, the methodology is very similar to smoothed regression techniques. The reader is referred to articles by Scott and Smith (1969) and Sedransk (1977) for a complete description of the technique.

5.3 *Implementation of methods*

Two interesting issues are raised by the consideration of allocation techniques: the question of how to form undercount estimates at all geographic levels, and how the allocation methods might affect the design of the undercount estimation measurement techniques. As mentioned in the previous section, synthetic estimates are guaranteed to sum to the totals used at higher levels of geography to make the estimates, but this is not true for regression or borrowed-strength estimates. If one wishes to retain the demographic or survey estimates at the levels for which they are formed, one can further weight the estimates at low levels of geography by the ratio of the estimates at higher levels to the sum of the estimates at lower levels. For the regression estimates, if they were regressed at the regional level and used to allocate down to the enumeration area, j, then the weight, W_i, for all estimates in region i would be formed as

$$W_i = \frac{\textit{regional estimate (ratio i) from PES}}{\sum_j \hat{U}_{ij}}$$

This weight applied to each of the $\hat{U}_{ij}$ as a revised enumeration area estimate would force the enumeration area estimates to sum to the regional estimate. The obvious advantage to this procedure is to make the lower level estimates consistent with estimates generated by survey or demographic methods.

This weighting technique can be used to make allocations to lower levels of geography, also. In fact, weights can be assigned to each individual and household to allow any tabulation necessary using weighted totals. These weights would be similar to noninterview adjustments made in sample surveys. Problems can arise using this technique (and others mentioned above) if attention is not paid to rounding errors, since they can lead to inconsistencies among tables. This problem can be reduced by careful attention to the programming of the tabulations, or the use of integer weights.

Use of integer weights has exactly the same effect as a "hot deck" imputation scheme. Integer weights would be randomly assigned to units within low levels of geography. For example, most housing units within an enumeration area would receive a weight of one, but one or more units would receive a weight of two to "double-count" some units or persons. Under hot deck imputation, the records for these units would simply be duplicated on the census data files. The most severe problem with this method is if there is an overcount measured for some demographic subgroups, this would imply assigning a zero weight to some units or individuals, effectively removing them from the census files. This would most likely be unacceptable politically since it implies throwing data away, although throwing data out at random is commonly used in experimental designs for other purposes.

Each of these methods has implications for the design of the data storage and tabulation systems employed by the census. An adjustment to the census has far-reaching implications for the various stages of the census, ultimately affecting even the computer systems used for processing the census. The adjustment has implications in the other direction also. If an adjustment is to be made, and regression or borrowed strength techniques are to be implemented, the evaluation program should be designed to allow the greatest flexibility for both measurement of the undercount and the use of broadly distributed small areas as units in the regressions. As with other phases of this program, there is a tradeoff between the need for generating estimates with low variance, implying larger areas represented in the sample estimates (assuming a PES-like study will be used for the regressions), versus having more units available for the regressions (the regression units being smaller and more diverse would each have a higher variance). Optimal designs would have to take account of the national level and an upper bound on the variance of estimates at regional levels to be used in the regressions.

6. *EVALUATION OF THE EVALUATION*

If the evaluation results are to be used to adjust the census, then an assessment of the quality of the evaluation is needed for two reasons. The first is that adjustment will have the effect of reducing the relative proportion of persons in some areas and therefore the resources, such as funding, allocated to that area by the national government. It is necessary therefore to confirm that the adjustment is based on the best data available, whether it is a PES, demographic analysis, or some combination. The second reason for an evaluation of the evaluation is to give a sense to the researcher of when the evaluation data are "good enough" to use in an adjustment, and if not, how much more research is needed. Applying the adjustment will have to be done fairly quickly after the census is complete, and so the research on the evaluation should take this into account. That is, the focus of part of the research should be on how to come to closure rapidly on deciding how to make the adjustment. A number of key decisions on how to make the adjustment cannot be made until the data are in hand, and evaluation research should be directed towards quickly determining which adjustment techniques will not be feasible.

7. *CONCLUSIONS*

Any type of adjustment introduced into the census will have far-reaching consequences. It will affect the way the census is conducted, how the results are published, the quality of the final data produced, and even noncensus operations conducted by the statistical office or other organizations which use census data for the design of surveys. The adjustment will be time consuming, costly, and require a significant amount of staff resources.

But an adjustment to the census may be judged to be necessary. If coverage for certain population groups is not up to the standards set by the statistical office, the adjustment may be of great benefit to the population groups disproportionately under-represented and to the economy as a whole.

A key conclusion that should be drawn from this chapter is that, although the adjustment may be expensive to the national statistical office, at the same time it may be necessary to reflect adequately the composition of the population being measured.

8. REFERENCES

Scott, Alastair, and Smith, Terence M.F. 1969. "Estimation in Multi-Stage Surveys." *Journal of the American Statistical Association* 64:830-840.

Sedransk, Joseph. 1977. "Sampling Problems in the Estimation of the Money Supply." *Journal of the American Statistical Association* 72:516-522.

Steinberg, Joseph, ed. 1979. *Synthetic Estimates for Small Areas: Statistical Workshop Papers and Discussion.* National Institute of Drug Abuse Research Monograph 24. Washington, D.C.: U.S. Department of Health, Education, and Welfare.

Chapter 7. PLANNING FOR CENSUS EVALUATION

1. INTRODUCTION

Planning for a census evaluation program has many facets in common with planning for the census itself. They include determining overall objectives; organizing, recruiting and training of staff; data collection, processing, analysis and publication; budgeting; developing and monitoring detailed schedules of operations; and dealing with relevant issues concerning publicity, confidentiality and legal mandates.

When field studies, such as a coverage study to measure an undercount or an interpenetration design to study systematic enumerator errors, are to constitute one or more components of the evaluation, the planning process takes on the additional requirements necessary for planning any moderately complex sample survey. In addition to the elements already mentioned for census planning, each of which is also necessary for any field study, there is a need for developing the evaluation questionnaire and the survey plan, designing the sample, pretesting, mapping, case-by-case matching (in a coverage study), and preparing the various estimates with requisite measures of precision.

Further, the use of administrative records and/or demographic techniques in the evaluation imposes still more planning requirements beyond those necessary for census-taking or for survey operations. With the use of demographic methods, allocation of (existing) staff becomes an important planning consideration. In addition, planning needs include determining the overall measurement methodology, including choosing among competing techniques; and acquiring, assembling and using various independent data files and data processing protocols.

Finally, even though the concept of independence is a guiding principle that pervades the design of an evaluation program, coordinated planning of the program cannot be independent of the census.

2. GENERAL REQUIREMENTS

In developing the overall plan for evaluating a population and housing census, there are many features of the process which are not unlike those for any large-scale data collection program, such as the census itself.

2.1 *Overall objectives*

Perhaps there is no greater obstruction to achieving a satisfactory evaluation program than failing to set forth, early enough, clear and unambiguous objectives to be served by the evaluation. This is true whether the objectives are modest or comprehensive. If the group in charge of overall census policy planning can settle upon what is wanted out of the evaluation, the task of designing how best to do it is made considerably easier for the technical evaluation staff. Too often, in the absence of well-defined and stated objectives, evaluation "planning" begins with unfocused discussions and arguments on how to do an evaluation rather than what to do. Though the feasibility of undertaking any evaluation effort has, of necessity, to be part of the planning process, deciding upon an appropriate, practical methodology for evaluation is, nevertheless, the second major activity to be developed. The first is to decide and agree upon what the evaluation is to find out. In ascertaining what an evaluation should encompass, there are many considerations, but the most important are scope, cost, resources and feasibility.

2.11 Scope.--The scope of a census evaluation program is highly dependent on the uses that are to be made of the census data itself. As a hypothetical example, one country may have a legislative requirement to obtain the "best estimate" of the population at the national level based on the census results. With prior evidence of undercounts in previous censuses, the census planning group could plausibly decide that the best strategy for fulfilling the legislative requirement would be to adjust the nationwide census total on the basis of a concurrent estimate of the undercount. The planning group might conclude, therefore, that the only objective of their evaluation program would be to estimate the undercount.

The above example specifies a simple objective, though not an easy one to implement. It is more likely that an evaluation program would serve multiple objectives. An appropriate scenario could be a country that wishes to:

(1) Estimate the census undercount at the national level, with a certain reliability;

(2) Estimate the census undercount at the regional, provincial, or other subnational level, at specified reliability levels that are equal, for example, by region but less than the specified reliability for the national estimate;

(3) Analyze evidence of age misreporting;

(4) Compare race/ethnicity estimates with independent data;

(5) Measure the extent of response variation on new census questions designed to tap internal migration;

(6) Study the correlated component effect (enumerator variance) to provide information to guide a decision on the possible use of self-administered questionnaires in a future census;

(7) Analyze the undercoverage estimates by various socioeconomic categories to the extent practicable, given reliability limitations, in order to identify population subgroups where improved coverage methods might be tried in the next census.

It seems clear that the scope of census evaluation may range broadly from having a single objective to having multiple objectives. Naturally, the more objectives there are, the greater the need for sound planning early in the census process. With multiple objectives, the planning function is the only way to safeguard against conflicting objectives or to provide the kind of oversight needed to ensure that other important objectives are not neglected.

2.12 Cost.--The first decision that the census planning group must make with respect to evaluation planning is whether there is to be a meaningful evaluation program or not. This is not strictly a policy or program consideration; cost is a determining factor. When one considers that a comprehensive evaluation program could account for 5 to 10 percent of the total census budget, certain choices must be made. For example, if sampling is used in the census enumeration to gather some of the questionnaire items, it is relevant to consider whether increasing the sample size (commensurate with the 5 to 10 percent evaluation allocation) to improve the reliability of estimates for small areas might not better serve the overall aims of the census than conducting a post-enumeration survey to measure undercoverage.

In most countries with scarce resources, this and other choices must be made in deciding whether to allocate funds for evaluation. It is necessary to make the appropriate trade-offs between evaluation and other competing activities, such as verification of coded items, nonresponse follow-up, verification of data entry, and a number of other desirable quality control efforts. It generally is felt in most countries today, however, that census evaluation is an integral part of census-taking, and so the ideal situation is for census planners to allocate some amount of funding for the conduct of a defensible evaluation program.

2.13 Resources.--Costs are clearly important, but the availability of other resources as well impinge upon the planning considerations in census evaluation. It may be necessary to increase the existing staff with several statisticians trained in demography, mathematical statistics, sampling, experimental design and survey methodology, as well as computer programmers with mathematical training. The need for specialized equipment, such as microcomputers for some of the technical analyses, has to be assessed in the initial stages of planning. Data processing requirements in particular can be troublesome if not properly planned; though the processing needs for an evaluation program are not likely to be as extensive as those for the main census, they are different and can be more complicated.

Hence, close examination of the overall resource requirements, by type and amount, is necessary to help shape the scope and scale of the evaluation program.

2.14 Feasibility.--Closely related to the resource considerations in planning for census evaluation is the issue of feasibility. For a country which may have never conducted a comprehensive census evaluation, it simply may not be feasible to plan a complex evaluation program requiring enormous technical resources or several stages of pretesting. On the other hand, many options for examining current census data for consistency and reasonableness are available to any country, so that technical feasibility should not be a particular hindrance for these methods of evaluation (see chapter 4).

2.2 *Organization, recruitment, and training*

The highest levels of management in the statistical agency responsible for the census must be fully committed to both the philosophy and the implementation of census evaluation. An almost certain way to doom an evaluation program is to try to carry one out without such commitment. Lack of institutional commitment can lead to damaging modifications or cutbacks in technical stages of the evaluation or even total abandonment of the program, under the combination of intense pressure and scarce resources which must be devoted to completing the main census.

Organizationally there should be designated a director of the census evaluation program who guides the totality of the effort. The evaluation director should, however, have no other responsibilities in connection with the main census operation. This person should be a member of the central census (policy) planning group and should report directly to the census director. Depending upon the scope of the evaluation program, the evaluation unit could be staffed in a variety of ways. A critical requirement, however, is the need for independence from the main census. The staff that are charged with carrying out the evaluation must be dedicated to that effort with no other operational responsibilities in the main census to the extent feasible. Independent evaluation is one of the standards of design, and this precept extends to the staff working on evaluation.

With regard to training, the selection criteria for choosing enumerators and the type of training given may differ considerably from that for the main census. Otherwise, there is nothing specialized about recruitment or training of either office staff or field personnel for census evaluation compared to the main census itself. It is only necessary to plan for training in the early stages of the census process in order to ensure that adequate staffing will be available to conduct the evaluation efficiently.

2.3 *Preparation of overall technical plan*

With the scope of the evaluation firmly in mind (see section 2.11), technical planning can begin. Here there are different levels of planning, especially if the evaluation scope is multifaceted with many diverse objectives. Just as it is the responsibility of the census planning group to ensure that the overall objectives are integrated and nonconflicting, that group also has an oversight responsibility for integration of the technical planning.

The responsibility for devising the detailed plans would be parceled out to several technical planning groups, each of which would be in charge of a particular component of the evaluation. For example, it might be reasonable and efficient to establish three technical groups, in charge, respectively, of (a) a post-enumeration survey, (b) a content reinterview check and (c) analysis using demographic techniques. Each group would develop its own technical plans, commensurate with the objectives mandated by the census planning group, and work somewhat independently of each other within specified guidelines for design. This is not to suggest that coordination of the planning functions carried on by each technical group is not essential (see section 5). These technical groups should, however, include some members of the main census operation to enhance communication and, consequently, the final product.

What are the basic elements of sound technical planning? First and foremost, the objectives of the evaluation have to be clear, unambiguous, and specific (this has already been discussed at length in section 2.1). Beyond that is the need for coordination with census operations, overall manageability, decisions on the choice of methodologies, and development of the evaluation design.

2.31 Relation to overall objectives.--As mentioned, the need for good, clear objectives is mandatory and has already been expounded. The reason this is so important to technical planning is that any statistical, technical or operational plan that is considered is much easier to formulate when the goals of the evaluation are well-understood by the design technicians.

2.32 Coordination with census operations.--Much evaluation work is dependent upon the census materials themselves. The availability of maps for a PES, the degree of tabulation detail of the census estimates for demographic analysis, and the timing of both the field work and data processing for the census are just a few examples of the many census activities that have to be taken into account by the evaluation specialists in planning.

It is apparent that thorough familiarization with the census plans and operational schedule is a prerequisite for evaluation planning. It is equally apparent that the evaluation plans must fit together with the census operations. This of course must be accomplished with minimum disruption to ongoing census activities.

2.33 Manageability.--The particular evaluation scheme that is proposed obviously must be manageable, in several ways. The complexity of the operational and measurement tasks set forth must be within the capability of the staff to carry them out, or else outside assistance must be planned. Timing considerations must be realistically planned with allowances for contingency operations. Anticipated use of data or materials from other ministries as input to the evaluation estimates must be researched in advance with respect to availability, relevance, timeliness and utility.

2.34 Choice of methodologies.--Initial decisions on the choice of suitable methodologies must be made by the technical planning group. Evaluation techniques are evolving with new and improved methods coming about rapidly. But different methods work in different countries, and what may be a useful evaluation tool in one census may be seen to be obsolete by the next. In many countries the choice of an evaluation measurement method is constrained severely, not by the lack of expertise or knowledge, but by the lack of suitable independent data that might be used for comparative purposes.

The evaluation planners thus have to choose methods and techniques that are robust and practical, yet germane and faithful to the objectives.

2.35 Development of design.--The design of an evaluation program can perhaps best be defined by the study methodology to be used. The study methodology is also referred to as the "research plan"; it entails the specification of the parameters of the design. These parameters include when and how the data will be collected, optimization determinants in balancing costs against variance and bias, the characteristics of the sample design (if sampling is used), the target population to be studied, the properties of estimators to be used, etc.

Where planning plays a key role in design development is in recognizing the need for modifications, which are inevitable, in the initial design parameters. Hence, planning for various design contingencies is probably necessary, or actually planning alternative designs may be required.

As an example, Procedure A for the PES (see chapter 2) cannot feasibly be used in a de jure census. The census planning group may be undecided about whether to use de jure or de facto methods and they may therefore plan to conduct some field tests before making the decision. The study design specialists in charge of the PES cannot "design" the use of Procedure A into the evaluation program until a decision is reached about a de jure versus a de facto enumeration. In this instance it would behoove the evaluation design group not only to plan for Procedure A conditionally (if it is otherwise deemed viable), but also to develop alternatives to Procedure A.

2.4 *Budgeting*

A comprehensive census evaluation can easily demand as much as 10 percent of the total budget for the census. Planning the budget for the census evaluation must be done at the same time planning the budget for the census takes place. When funds for the census are legislatively authorized or sought from donor organizations, they must include, specifically, a line item for the evaluation component. Detailed budgets will have to be developed by each staff responsible for the various evaluation modules (e.g., the PES program, the demographic analysis program, etc.) which make up the overall evaluation effort.

2.5 *Development and monitoring of schedules*

Planning to develop detailed schedules of all operations in the evaluation program is recommended. For a multifaceted effort, a separate detailed schedule for each evaluation study or component is necessary.

It is recommended that flowcharting of major operations and activities precede the preparation of detailed schedules. Again, the more comprehensive or complex the evaluation, the more this is needed. Ideally the flowchart(s) would show the integration and relationship of operations among the separate evaluation components,

and the relevant nections with main census operations (for example, how matching PES cases fits with census processing).

With the existence of good schedules, evaluation managers can monitor progress more easily, and in doing so, they can also see where alterations in plans must be made because of unforeseen snags.

2.6 *Publicity, confidentiality, and legal mandate*

In most countries a periodic population census is legislatively authorized. And, also by law, response to the census questionnaire is mandatory. It is important that there also be a legislative basis for conducting a census evaluation.

When second interviews are taken with a sample of respondents for evaluation purposes, whether with the same questionnaire or a different one, it is critical that respondents understand why. Enumerators must be prepared to give a brief explanation of the reason for the second interview and how it is to be used. In this connection advance publicity of the fact that some households, to be chosen strictly at random, will be asked to respond to a second census-related evaluation questionnaire is useful to foster cooperation and high response rates.

When there is a legal mandate for the evaluation it is easier to plan both an appropriate publicity campaign and the content of prepared statements that enumerators can use to persuade reluctant respondents. Of course, if response to the evaluation questionnaire is mandatory, this requirement must be made known to respondents.

Related to the legal aspects of data collection is the issue of confidentiality. In countries where individual responses to census questions are confidential and used only for statistical aggregation, the evaluation responses should also be treated confidentially. Respondents must be informed of this. Reinterview surveys are sometimes particularly difficult because some respondents may feel that they are being checked on due to something being "wrong" with their responses in the main census questionnaire. This sensitive problem can be lessened if the enumerator enumerator can emphasize the random manner in which the respondent was chosen, the confidential nature of the questionnaire, and that the results are to be used only for statistical purposes.

The confidentiality issue on the use of administrative records is also important in planning the evaluation program. For example, if civil registration lists are to be used to do a record match to estimate the number of aliens, it would be necessary to plan what exchange of information between the census agency and the civil registration agency would be kept confidential, and how it would be done.

3. *SPECIAL REQUIREMENTS FOR FIELD STUDIES*

There are a host of concerns related to planning that are generally applicable to census evaluation as discussed in section 2. Depending upon the specific nature of the evaluation program, there are additional concerns. This section deals with those specific planning issues for an evaluation program that includes field studies or investigations.

3.1 *Questionnaire design and survey plan*

The planning which is necessary for the design of field questionnaires and, in general, the development of the overall evaluation survey is not conceptually different from that for any sample survey. As such, there is the need to plan the questionnaire content, wording and layout; the timing and distribution of the field work; the timing and method of processing the completed questionnaires; etc.

Where planning for questionnaire design/survey operations is distinct for census evaluation is in the necessity for complete integration or coordination with the main census. Questionnaire content for evaluation is unavoidably tied to a subset of census content, although the exact wording and design are not. For example, a reinterview study might be designed to assess response variation on census reporting of single years of age; the measurement instrument, or questionnaire, might be radically different in approach, sequencing, or wording than the census questionnaire, though both may contain the same substantive content for age measurement.

Thus in planning the evaluation survey, planners must be keenly aware of the need for sufficient lead time to develop the instruments. It is particularly likely, for example, that pretesting various versions of evaluation questionnaires will be necessary when the proposed instrument is otherwise untried. Of equal importance is the need for evaluation planners to take heed of how changes in census plans can affect the evaluation plan. For example, the census planning group may conduct a pretest of several questions on disability before deciding whether to include those questions on the final questionnaire. The evaluation group must, accordingly, coordinate its plans with the census planning group if it intends to evaluate the quality of response on disability.

3.2 *Sample design*

The keys to proper planning for the sample design of an evaluation survey include (a) determining the domains of study, (b) determining the source of sampling frame, and (c) seeking independence of design and implementation.

3.21 Domains.--Planning should be done, at the outset, on what the domains of study should be in the evaluation survey. Domains are the geographic areas or, in some instances, the demographic or socioeconomic subgroups for which separate estimates are needed with certain, specified degrees of precision. Determining the definitions of the domains follows logically from setting the evaluation survey objectives (see section 2.1). Thus, planning to specify the study domains clearly is a prerequisite for the technical staff in charge of designing the sample. With the domains clearly delineated, the way is paved for working out the design issues, such as determining overall sample size, deciding on the number and make-up of the stages of sample selection, and determining how to build in sampling efficiencies such as clustering or stratification.

3.22 Sampling frame.--A very basic planning issue is deciding on what source materials are to be used to develop the sampling frame for the evaluation survey. The key is in whether the frame is to be developed from the census being evaluated, or from some other source. For a coverage evaluation survey, a source apart from the current census is mandatory in order to fulfill the independent measurement requirement. On the other hand, a content reinterview survey might plausibly make use of either the current census listings or a current household survey sample as the sampling frame.

In either case, important timing considerations must be taken into account in planning. If the current census is to constitute the frame, then planning must proceed on how and when to get access to the census listing books, maps, enumerator assignments, or other materials for purposes of constructing the frame. Part of the planning effort must deal with the problem of minimizing interruption to the flow of census documents for census processing, while at the same time seeking to get the evaluation survey frame developed in a timely manner. This can be a crucial problem since the evaluation survey must usually follow the main census enumeration within a very short period of time.

When an independent source, such as an area-based frame for a coverage study, is used for the sampling frame, planning for the development of that frame can proceed without concern about the main census materials as such. However, planning may have to begin sooner than that for the census-based frame, especially if the independent source must entail a considerable amount of development in its own right.

3.23 Independence.--Throughout this manual, independence has been stressed as one of the rules of census evaluation. Independence is especially important for the sample design. Planners should make provisions for a separate staff, if possible, to be responsible for designing the evaluation survey sample, that is, a staff separate from the group which would have responsibility for the use of sampling and its design in the main census. As noted, an independent area-based frame is a must for a coverage evaluation survey. In the field, the individuals responsible for listing and enumerating sample households should be independent of those working the same geographic areas in the census.

3.3 *Mapping*

The use of maps in the main census requires an enormous amount of planning. Planning for the use of maps in census evaluation surveys is no less important, though less extensive. For example, the use of segment maps may be the only cartographic requirement in a PES, and the scale of a PES is, of course, much smaller than the census. It is not possible to conduct a PES, in fact, without maps, since the unit-by-unit coverage has to be judged in area segments with well-defined boundaries. Appropriate planning for the timing of the preparation of the evaluation survey maps is, therefore, an important responsibility of the technical planning group.

3.4 *Pretesting*

It is likely that a certain amount of pretesting will be done in preparing for the main census, including small-scale field tests, pilot testing, and perhaps a full-scale "dress rehearsal." Just as likely is the possibility that evaluation survey pretesting will be done.

With a content reinterview study, the questionnaire can be different from the main census questionnaire (see section 3.1). In this case, it is essential that some pretesting be done to establish feasibility, determine necessary modifications, etc. For a PES or interpenetrated study design, the survey methods are more complex, and so pretesting of the overall field methodology probably is necessary before engaging in the full-scale operation. For a PES, it is also particularly important to test the overall workability of the proposed matching procedures.

In general, it is vital that countries which have never undertaken evaluation surveys plan for appropriate pretests. Even those countries which have experience in evaluation surveys would benefit from planning certain pretests in order to refine their measurement procedures.

3.5 *Matching*

Matching of individual records between the evaluation survey and the main census is one of the distinguishing features of a PES or content reinterview program. It is also one of the most difficult; if not done well, the evaluation survey results can be rendered useless. Devoting sufficient attention to planning the intricate requirements for the entire matching operation (whether exclusively manual or computer-assisted), establishing matching rules and reconciliation procedures, integrating the flow of records between census and evaluation, etc., is one of the most demanding and critical responsibilities of the technical planning group.

3.6 *Preparation of estimates and measures of precision*

Since an evaluation survey, whatever its design or purpose, is based on a complex sample plan and enumeration, then the preparation of the evaluation estimates will likely be somewhat more complicated (and will certainly be of a different nature) than the main census estimates. Even if sampling is used in the main census, it often is done in such a way that the resulting estimation procedure is straightforward. Moreover, various measures of reliability, such as sampling error and certain bias components of the total mean square error associated with the evaluation estimates themselves, will also be computed. It is necessary, therefore, to plan for these added complications, with respect to their impact on such matters as the need for technical assistance, additional timing to carry the assistance out, computer programming requirements, and presentation of results.

4. *SPECIAL REQUIREMENTS OF DEMOGRAPHIC TECHNIQUES*

The preceding two main sections of this chapter have discussed the planning needs (a) generally applicable for census evaluation and (b) those necessary especially for field studies. This section takes up planning needs which are especially pertinent to demographic techniques of evaluation, those that do not rely upon original data collection in the field.

4.1 *Staff development*

In many respects it can be argued that demographic analysis for the evaluation of census data is a continuing process, extending in unbroken activity from one census to the next. Indeed many countries retain a small professional staff of demographers and statisticians who are employed full time for the purpose of analyzing census data, including assessing its quality.

Planning the demographic components of census evaluation, therefore, may not call for substantial augmentation of staff but rather a shift of focus for the existing staff from their regular analysis tasks to the special-purpose tasks of designing and executing the evaluation program. In countries where little or no demographic evaluation takes place, the census planning group must, in the process of assessing its plans for the demographic evaluation component, decide what staffing is needed, so that budgeting and recruiting for evaluation can be done.

4.2 *Overall measurement methodology*

There is a wide array of demographic methods that can be used to evaluate census estimates (see chapter 4). The choice of appropriate techniques must take into account the cost, availability, and quality of independent data, the complexity of applying one technique versus another, and, of course, considerations of primary versus secondary objectives to be served by the evaluation.

So, development of an overall plan for the demographic evaluation methodology is necessary to achieve a cohesive and integrated evaluation; one that meets the proposed objectives but is within the limits or constraints of staff capability and resources. Moreover, the demographic evaluation plan should be developed along with the plans for other aspects of the total evaluation package, coverage, reinterview, etc., to achieve a properly coordinated and comprehensive census evaluation program.

4.3 *Obtaining independent data sources*

Many data sources can be used in the various demographic evaluation methods. One of the most obvious sources, which is independent from the current census, is data from previous censuses. Another may be current survey data from the country's intercensal survey program. Still more sources might be vital statistics records, civil registration records, income tax records, or social insurance records.

Plans must be made for accessing each data source that will be needed, acquiring the data, manipulating it to provide the necessary comparative estimates, and preparing some fairly detailed documentation. To the extent that some data sources may be jurisdictionally housed in other ministries, early planning is necessary to allow for formalities, questions of confidentiality, technical difficulties of data conversion to the proper computer format, etc.

5. *COORDINATED PLANNING*

We have stressed throughout this chapter that good evaluation is served by independence of effort. That is, a sound census evaluation program is best accomplished when the evaluation design and implementation are carried out by a dedicated staff whose duties are independent of main census operations. This is not always possible in countries where resources are limited; but to the extent practical, it is a goal to strive toward. Certain technical requirements, however, such as not allowing the census field staff to know which area segments are selected for the PES, exemplify the necessity for separation of effort.

In spite of the need for independence in the technical facets, however, there is one aspect of census evaluation where independence of effort would be detrimental; that is in the area of planning. Planning, at all levels, must be coordinated and interdependent. The central census planning committee must incorporate the evaluation plans as part of the overall census program; this committee must therefore have representation from the evaluation group, probably the census evaluation director. Further, specific components of an evaluation program require coordinated planning, with each other (e.g., PES, interpenetration study, reinterview) and with the main census. It is essential therefore that the managers in charge of those specific studies work together for integrated planning.

APPENDIX

THE THEORY AND METHODS OF THE STATISTICAL EVALUATION OF CONTENT ERROR

1. INTRODUCTION

This appendix describes the statistical theory for the methods used in chapter 3 in the evaluation of census content error. This theoretical development assumes familiarity with the concepts of expected value, variance, and general statistical modeling techniques. It is included for those readers with a good knowledge of basic statistics who require a deeper understanding of the concepts and methods of statistical evaluation in order to design, analyze, and interpret the results of evaluation studies.

2. THE GENERAL MODEL AND FORMULATION

In this section, a general model for investigating response error in both sample surveys and censuses will be described. Sample surveys are included in our discussion because:

(1) Most census evaluation studies are conducted on a sample basis.

(2) The census itself may incorporate a sample component for some items.

(3) Some commonly used statistical measures of content error were originally developed for sample surveys and are, therefore, best described from this perspective.

We shall focus on one particular item on the census questionnaire in our development. For our purposes, the term census population, or population, is the collection of all units targeted in the census for this item. For example, the units may be persons, households, or housing units possessing some screening characteristic.

Let N denote the size of the census population. Let n denote the size of a sample drawn from the population for some purpose--e.g., an evaluation survey, a census survey component, or otherwise.

Denote by the subscript j, the j^{th} unit in the sample, by y_j the final recorded response for the j^{th} sample unit, by μ_j the true value of unit j for the item, and by e_j the deviation of y_j from μ_j, that is the error in unit j's response. We have just described a very simple response error model, namely,

$$y_j = \mu_j + e_j \tag{2.1}$$

or, in words,

$$\text{Recorded value} = \text{True value} + \text{Error}$$

The statistical evaluation of census content error strives to determine the impact of the errors, e_j, on census totals, proportions, and means. It assumes that the e_j, are random variables and attempts to estimate their means, variances, and covariances, since these components comprise the total error in census statistics.

Let Y denote the census total for some characteristics. The total error in Y is measured by the mean square error *(MSE)* which is the bias of Y squared plus the variance of Y. That is

$$MSE(Y) = [Bias(Y)]^2 + Variance(Y) \tag{2.2}$$

For example, the census total, Y, is given by

$$Y = \sum_{j=1}^{N} y_j$$

which by model (2.1) is

$$= \sum_{j=1}^{N} \mu_j + \sum_{j=1}^{N} e_j$$

Denote the expected value of e_j with respect to its underlying statistical distribution by

$$E(e_j) = B_j \tag{2.3}$$

and denote $\frac{1}{N}\sum_{j=1}^{N} B_j$, the mean of B_j over all population units by B, called the census <u>response bias</u>. Thus,

$$(2.4) \qquad E(Y) = \sum_{j=1}^{N} \mu_j + N \cdot B$$

or E(Census total) = True population total + Bias

Formula (2.4) also holds for sample surveys. Let $\hat{Y}$ denote the usual unbiased estimator of Y, where the computational form of $\hat{Y}$ is determined by the sample design. Then

$$(2.5) \qquad E(\hat{Y}) = \sum_{j=1}^{N} \mu_j + N \cdot B$$

where the expected value operator, E, now also includes the sampling distribution as well as the nonsampling error distribution(s).

The response bias objective of content reinterview defined in section 2.2 of chapter 3 was concerned with estimating B. As we see from (2.4), this is possible only if an estimator of the true total, $\sum_{1}^{N} \mu_j$, is available.

Now, consider the second part of the mean square error of Y(or $\hat{Y}$); viz., the variance. Let $V(Y)$ denote the total variance of the census total, Y. By model (2.1), we have

$$\begin{aligned} V(Y) &= V\left(\sum_{j=1}^{N} y_j\right) \\ &= V\left(\sum_{1}^{N} \mu_j + \sum_{1}^{N} e_j\right) \\ &= V\left(\sum_{1}^{N} e_j\right) \end{aligned}$$

since $\sum_{1}^{N} \mu_j$ is the population total, a constant. Finally, $V\left(\sum_{1}^{N} e_j\right)$ and, hence, $V(Y)$ can be written as

$$(2.6) \qquad \sum_{1}^{N} V(e_j) + \sum_{j \neq}^{N} \sum_{j'}^{N} Cov(y_j, y_{j'})$$

where Cov denotes the covariance. The first term, divided by N, is known as the <u>simple response variance</u> (SRV). It is the average variance of responses from the same unit to the same question over repeated inquiries.

The second term in (2.6), the covariance term, arises as a result of external factors in the census that cause similar errors in responses from different units. One of the major contributors to this term, as we mentioned in section 1.3 of chapter 3, is the census enumerator. Other factors might be supervisors, trainers, and coders; however, we will be primarily interested in enumerator covariances. Therefore, assume that $Cov(e_j, e_{j'}) = 0$, if units j and j' are in two different enumerator assignments. If j and j' are two units in the same enumerator assignment, $Cov(e_j, e_{j'})$ may not be zero. This covariance is called the correlated component of enumerator variance denoted by CC.

Suppose there are k enumerators for the census labelled $i=1, \ldots, k$. Let m_i denote the number of units in the i^{th} enumerator's assignment. The covariance term in (2.6) can be rewritten as

$$(2.7) \qquad \sum_{i=1}^{k} m_i(m_i - 1)CC$$

since there are $m_i(m_i-1)$ pairs of units in an assignment of size m_i each having covariance CC.

We can now summarize these results in one formula for the mean square error of a census total Y:

$$(2.8) \quad MSE(Y) = N^2B^2 + \left(\sum_{1}^{k} m_i^2 - N\right)CC + N \cdot SRV$$

Thus, there are three components that make up the total error in a census total: the bias (B), the simple response variance (SRV) and the correlated component (CC). In the remaining sections of this appendix, we shall discuss (a) the estimation of these components and (b) two specific models that aid in the interpretation of the components.

It is interesting to view (2.8) assuming that each enumerator assignment is approximately the same size, i.e., $m_i = m$, say for all enumerators $i=1, \ldots, k$. Formula (2.8) is, thus,

$$(2.9) \quad MSE(Y) = N^2B^2 + N\cdot(m - 1)CC + N\cdot SRV$$

which emphasizes that the impact of CC on the total census error depends upon the size of the enumerator assignment or, equivalently, the number of enumerators employed for the census. The last term in (2.9) is at a maximum when $m = N$ (or $k = 1$); that is, when only one enumerator collects information from all N units. The impact of CC is at a minimum when $m = 1$ $(k = N)$ or when every population unit is enumerated by a different enumerator. Obviously, these two extremes can never be achieved in practice--or can they? For example, one enumerator may collect the census information for all persons in a small village. Therefore, the enumerator has maximal impact on that village's statistics. On the other hand, the census questionnaire may be filled entirely by respondents without the assistance of the enumerator. Thus, each respondent is his or her own enumerator $(k = N)$ with minimal enumerator impact on census error. In most cases, the number of enumerators employed in the census is somewhere between these two extremes. However, this number should be determined, in part, with consideration of the impact of the correlated component of enumerator variance on the statistics.

Finally, let us consider the variance of an estimated total based upon a sample survey. In addition to the variability in the estimator due to response errors, we must also consider the variability arising from sampling, i.e., the sampling variance. However, the sampling variance depends upon the sample design implemented for the survey, and general formulas which apply for all sample designs are conceptually difficult. Therefore, to illustrate the basic concepts, we shall assume simple random sampling.

A general formula for the variance of $\hat{Y} = N\bar{y}$, where $\bar{y} = \sum_{j=1}^{n} y_j/n$ is the mean of a simple random sample of size n, will now be given assuming, for simplicity, equal interviewer assignment sizes $(m_i = m)$. The variance of $\hat{Y}$ is $N^2V(\bar{y})$ where $V(\bar{y})$ is the variance of the sample mean $\bar{y}$ given by

$$(2.10) \quad V(\bar{y}) = \left(1 - \frac{n-1}{N-1}\right)\frac{SV}{n} + \frac{m-1}{m}\frac{CC}{k} + \frac{SRV}{n}$$

and

$SV = \underset{j}{V}E(y_j|_j)$ is called the sampling variance. It is the variance over the census population of the $\mu_j + B_j$ where B_j is the unit response bias defined in (2.3).

$CC = Cov(y_j, y_{j'}) - Cov(y_a, y_b)$ is a more general definition of the correlated component previously defined. $Cov(y_j, y_{j'})$ is the covariance between any responses for two units in the same interviewer's assignment and $Cov(y_a, y_b)$ is the covariance between any responses for two units in different interviewer assignments. Therefore, CC is the covariance that is specific to units within an interviewer's assignment after eliminating the covariance that is common to all units without regard to interviewer assignments.

$SRV = \underset{j}{E}V(y_j|_j)$ or since $V(\mu_j|_j) = 0$,

$= \underset{j}{E}V(e_j|_j)$, the simple response variance for sample surveys. Note that $\underset{j}{E}V(y_j|_j) = \frac{1}{N}\sum_{1}^{N} V(e_j)$ which was defined previously in (2.6).

Formula (2.10) applies under very general conditions as will be discussed in the next section. It is easy to verify that (2.10) and $V(Y)$ given in (2.9) are identical when $n = N$, a complete census. It ignores the correlation between errors in different enumerator assignments that may arise in a survey as the result of trainers, supervisors, other survey personnel, and respondents. However, experience suggests that these correlations are typically small relative to interviewer correlations. In addition, it does not take into account the correlation between the error made for a particular unit in an interviewer's assignment and the true values of the other units in an assignment. Although little is known about its magnitude, this covariance, referred to as the interaction covariance, may be substantial. An example of this occurred in the 1950 U.S. census. Enumerators were to provide estimates based on observation of the cost of housing units in their assignments. Housing units in assignments with a large proportion of high-cost housing tended to be overestimated. Conversely, a high proportion of low-cost housing in the assignments tended to bias the estimates downward.

An alternative form for the variance of $\bar{y}$ is often encountered in the literature. This form is expressed in terms of the intra-interviewer correlation coefficient denoted by

$$(2.11) \qquad \rho_y = \frac{CC}{SV + SRV}$$

$$(2.12) \qquad \doteq \frac{Cov(y_j, y_{j'})}{V(y_j)}$$

where j and j' are units in the same interviewer assignment, and the approximation is due to ignoring the between interviewer covariance in the definition of CC. Note that $SV + SRV = V(y_j)$ by the well-known identity $V(y_j) = \underset{j}{VE}(y_j|_j) + \underset{j}{EV}(y_j|_j)$. With this definition, $V(\bar{y})$ can be expressed as

$$(2.13) \qquad \frac{SV + SRV}{n}\left[1 + (m - 1)\rho_y\right] - \frac{SV}{N - 1}\,\frac{n - 1}{n}$$

The last term is due to the finite population correction. It is usually dropped for surveys where n/N is very small, because then the first term dominates the variance. Under certain conditions, the correlation coefficient, ρ_y, can be estimated from either census or survey data as will be discussed in section 4. These estimates of ρ_y can then be compared with other estimates published (in the literature) for census and survey data to gauge the magnitude of the enumerator effect. Figure 3-10 in chapter 3 provided general rules for ρ_y obtained from just such an analysis.

Another measure of response error commonly used in content error evaluation is the index of inconsistency denoted by I and given by

$$(2.14) \qquad I = \frac{SRV}{SV + SRV}$$

Whereas ρ_y measures the magnitude of CC relative to $V(y_j)$, I measures the magnitude of the simple response variance to the total variance, $V(y_j)$. Both ρ_y and I will be discussed in detail later.

The next section focuses on the interpretation of the components CC and SRV (or ρ_y and I) for two types of data often encountered in censuses and surveys: quantitative data and qualitative data. Quantitative data are measured on a continuous scale. Examples are income, age, years of education, etc., when these are reported as a single number. Qualitative data are reported by categories. Examples are sex, race, religion, income group, etc. The nature of the error, e_j, differs by type of data, as we shall see. For example, for income (reported as a single number), e_j may theoretically take on any value within some interval about 0. For sex reported as male = 0 and female = 1, however, e_j may only take on three values: -1, 0, or 1. Thus, the structure of the error e_j depends on the type of data we are collecting.

Two models which describe the structure of the response error will be examined in the next section. Through these models, we can better interpret the meaning of a high or low value of *CC* or *SRV*. Additionally, these models suggest statistical measures or indices that help us to summarize the results of evaluation studies and to compare these results with other studies both within our own country and in other countries.

3. ERROR STRUCTURES AND THEIR USE

In this section, the structures for the error e_j which are appropriate for both quantitative and qualitative data will be examined using the model $y_j = \mu_j + e_j$. Through these alternative structures, the way in which response error affects $V(\bar{y})$ will be demonstrated.

3.1 *Quantitative data*

As a notational convenience, we shall replace the single subscript j denoting the unit in the sample or the census by the double subscript (i,j) denoting the j^{th} unit in the i^{th} interviewer's assignment. Thus, the e_j, $j=1, \ldots, n$, become e_{ij}, $i=1, \ldots, k$ and $j=1, \ldots, m_i$.

The structure of e_{ij} for quantitative data is assumed to be

$$e_{ij} = b_i + \varepsilon_{ij} \tag{3.1.1}$$

where b_i is the systematic error associated with the i^{th} enumerator, and ε_{ij} is the variable error associated with unit (i,j). ε_{ij} is a combination of the variable errors of all other sources of error in the survey. Thus, model (2.1) may be written as

$$y_{ij} = \mu_{ij} + b_i + \varepsilon_{ij} \tag{3.1.2}$$

This model may be recognized as the usual "analysis of variance" model.

The model assumes the following:

(1) The enumerator variables $b_i, \ldots, b_k$ are a random sample from an infinite population of enumerator variables.

(2) $E(b_i) = B_b$ and $V(b_i) = \sigma_b^2$

(3) $E(\varepsilon_{ij}) = B_\varepsilon$ and $V(\varepsilon_{ij}) = \sigma_\varepsilon^2$

(4) b_i and ε_{ij} are mutually independent

(5) For sample surveys, model (3.1.2) pertains to unit (i,j) in the sample, and the sample value μ_{ij} is independent of b_i and ε_{ij}.

Under this model we will derive expressions for B, $V(\bar{y})$, ρ_y and I. Sample surveys will first be considered. These results will be re-examined for use in censuses.

Assume simple random sampling for the sample survey. The usual "unbiased" estimator of the population mean is $\bar{y} = \sum_i \sum_j y_{ij}/n$, the sample mean. By assumptions (2) and (3) and model (3.1.2), the expected value of $\bar{y}$ is

$$E(\bar{y}) = E\left[\frac{\sum_i \sum_j (\mu_{ij} + b_i + \varepsilon_{ij})}{n}\right] = \mu + B_b + B_\varepsilon \tag{3.1.3}$$

where μ is the true population mean. Therefore, the bias in $\bar{y}$ is $B = B_b + B_\varepsilon$. Part of the bias (B_b) is due to interviewers, and part (B_ε) is due to respondents and other sources of random error.

For simplicity, assume $m_i = m$ for all interviewers. Using the definitions presented in section 2, it is possible to show that

$SV = \sigma_\mu^2 = \sum_{j=1}^{N} (\mu_j - \bar{\mu})^2/N$, the population variance for the characteristic.

$CC = \sigma_b^2$, the population variance for interviewers, or the interviewer variance.

$SRV = \sigma_b^2 + \sigma_\varepsilon^2$, the interviewer variance component plus the random error variance component.

Now, using the general variance formula for surveys, (2.10), we have

$$(3.1.4) \qquad V(\bar{y}) = \left(1 - \frac{n-1}{N-1}\right) \frac{\sigma_\mu^2}{n} + \frac{\sigma_b^2}{k} + \frac{\sigma_\varepsilon^2}{n}$$

This equation demonstrates the impact of the correlated response error on the variance of $\bar{y}$. It is evident that the variance due to correlated errors does not decrease as the sample size increases, but only as the number of interviewers increases. Hence, when n is large, $V(\bar{y})$ may be dominated by the term σ_b^2/k.

For this model, the interpretation of $CC = 0$ or equivalently, $\sigma_b^2 = 0$, is that $b_i = B_b$ for all interviewers (i.e., there is no variability among interviewer biases). Since this suggests that the interviewers are performing consistently, one might interpret this as an indication that the operation is in control and is producing quality results. During the 1950 U.S. census, an experiment was conducted to estimate the component σ_b^2 for census enumerators. This study showed that for small area statistics σ_b^2, the enumerator contribution to the total variance was considerable. This led to the adoption of the method of self-enumeration beginning with the 1960 census (see chapter 3, section 4.3). For self-enumeration surveys, the correlated component, $\sigma_{b'}^2$, may be interpreted as the variance of the respondent error variables, b_i'. That is, each respondent may be considered as his/her own interviewer so that the correlated component has the divisor n. Thus if $(\sigma_{b'}^2 + \sigma_\varepsilon^2)/n$ for self-enumeration is smaller than $\sigma_b^2/k + \sigma_\varepsilon^2/n$ for the enumerator assisted census, $V(\bar{y})$ can be decreased if self-enumeration is adopted.

Using (2.13), $V(\bar{y})$ can also be written ignoring the finite population correction as

$$(3.1.5) \qquad V(\bar{y}) = \frac{\sigma_\mu^2 + \sigma_b^2 + \sigma_\varepsilon^2}{n} \left(1 + (m-1)\rho_y\right)$$

where, for the present model, ρ_y, with general form given by (2.11), has the specific form

$$(3.1.6) \qquad \rho_y = \frac{\sigma_b^2}{\sigma_\mu^2 + \sigma_b^2 + \sigma_\varepsilon^2}$$

When the sampling fraction n/N is small, $Cov(y_{ij}, y_{ij'})$ for two units (i,j) and (i,j') in the same interviewer assignment is σ_b^2. Also, $V(y_{ij}) = \sigma_\mu^2 + \sigma_b^2 + \sigma_\varepsilon^2$ so that (3.1.6) can be written as

$$(3.1.7) \qquad \rho_y = \frac{Cov(y_{ij}, y_{ij'})}{V(y_{ij})}$$

which is the correlation coefficient between units in the same interviewer assignment. Thus, in analogy to the intra-cluster correlation coefficient defined in cluster sampling, ρ_y has been called the intra-interviewer assignment correlation coefficient. Note, however, that ρ_y is also the ratio of interviewer variance, σ_b^2, to the total variance of the recorded value y_{ij}.

Using (3.1.5) the effect of small correlations between the units introduced by interviewers can be illustrated. Even a small ρ_y may make a major contribution to $V(\bar{y})$ since it is multiplied by roughly the size of an interviewer's assignment. This is discussed below in example 3.1.

Using (2.14), the index of inconsistency for quantitative data has the form

$$(3.1.8) \qquad I = \frac{\sigma_b^2 + \sigma_\varepsilon^2}{\sigma_\mu^2 + \sigma_b^2 + \sigma_\varepsilon^2}$$

Thus, I is the ratio of the total response variance, $V(b_i + \varepsilon_{ij})$, to the total variance, $V(y_{ij})$.

Example 3.1--Consider the impact of interviewer variance in a national fertility survey. Suppose that, for the characteristic "number of children ever born to the respondent," the parameters in (3.1.5) are $\sigma_\mu^2 + \sigma_b^2 + \sigma_\varepsilon^2 = 6.0$, $\rho_y = .01$ and the sample size, n, is $1,000$. If 100 interviewers were hired for the survey, each interviewing approximately the same number of respondents $(m=10)$, then the value of $(1 + (m-1)\rho_y)$ is 1.1. If 10 interviewers were hired, then $m=100$ and $(1 + (m-1)\rho_y)$ is 2.0. Therefore, if only 10 interviewers are employed, the variance of $\bar{y}$ is 1.82 times greater than if 100 interviewers are employed, assuming the interviewer variance is the same in both cases. Just as sampling variance decreases as the sample size increases, interviewer variance decreases as the number of interviewers increases.

It can be shown that, for a complete census, the expressions for *SV*, *CC*, and *SRV* are identical to those given for sample surveys. The variance of $\bar{y} = \bar{Y}$ can be obtained directly from (3.1.4) setting n equal to N. For a complete census

$$V(\bar{Y}) = \frac{\sigma_b^2}{k} + \frac{\sigma_\varepsilon^2}{N} \tag{3.1.9}$$

Thus, the total variance of a census mean contains a component for enumerator variance, which is divided by the number of enumerators in the census, and a component for random variance which is divided by the size of the census population. The interpretation of ρ_y and I is essentially the same for censuses as it is for surveys.

Note that the form of the variance given in (3.1.5) is not appropriate for determining the impact of enumerator variance of the total variance of $\bar{Y}$, since the finite population term shown in equation (2.13) was ignored. Therefore, for censuses, equation (3.1.9) is the appropriate one.

3.2 *Qualitative data*

For categorical data, the additive component structure for enumerator or interviewer error given in the last section is not appropriate, since the true value μ_{ij} takes only the values 0 and 1. A more appropriate model, provided in (Bailar and Biemer 1984), is as follows.

Let $\mu_{ij} = 1$ if unit (i,j) belongs to some class, say, C, and $\mu_{ij} = 0$ otherwise. Now we must consider errors contributed by an interviewer that result in the misclassification of a unit. For example, suppose we are interested in evaluating the interviewers in a national unemployment survey. Suppose for one of the items on the questionnaire the interviewers must classify an individual as being a member of the labor force or not a member. Given that a person is not a member of the labor force, interviewer i has a probability, say ϕ_i, of misclassifying the individual as being in the labor force. And, for a person who is in the labor force, the interviewer misclassifies at a rate θ_i. Each interviewer's rates of misclassification, ϕ_i and θ_i, are a function of the interviewer's understanding of the concepts, personal biases, thoroughness during the interview, etc., and, thus, ϕ_i and θ_i may vary from interviewer to interviewer.

Therefore, θ_i denotes the probability that interviewer i causes a unit in class C to be misclassified as not in class C (for example, someone unemployed is classified as being employed) and ϕ_i denotes the probability that interviewer i causes a unit not in class C to be misclassified into class C (i.e., someone employed is classified as being unemployed). Mathematically, this is written as

$$\theta_i = Pr(y_{ij} = 0 \mid \mu_{ij} = 1)$$

and

$$\phi_i = Pr(y_{ij} = 1 \mid \mu_{ij} = 0)$$

Now e_{ij} in the general model (2.1) can take on three possible values, -1, 0, or 1.

Value of μ_{ij}	Value of y_{ij}: 0	1
0	$e_{ij} = 0$	$e_{ij} = 1$
1	$e_{ij} = -1$	$e_{ij} = 0$

In terms of the general model (2.1), if $\mu_{ij} = 1$, then $e_{ij} = -1$ or 0 with probabilities θ_i and $1 - \theta_i$, respectively; and if $\mu_{ij} = 0$, then $e_{ij} = 1$ or 0 with probabilities ϕ_i and $1 - \phi_i$, respectively. By allowing the "probability of a false negative" θ_i, and the "probability of a false positive," ϕ_i, to depend upon the interviewers, we are able to investigate the impact of interviewer variability on the estimates of proportions.

It is assumed that the misclassification probabilities, θ_i and ϕ_i, $i=1, \ldots, k$, corresponding to the k interviewers for the survey, constitute a random sample from an infinite population of interviewer misclassification probabilities with means (θ, ϕ), variances $(\sigma^2_\theta, \sigma^2_\phi)$, and covariance $\sigma_{\theta\phi}$.

It is now possible to derive new expressions for the bias and variance components of total response error. As before, we shall first focus on sample surveys for the interpretations of the components; however, the interpretations for censuses are identical.

Assume simple random sampling for the sample survey as before. Then, it can be shown that

(3.2.1) $$E(p) = \mu - P\theta + Q\phi$$

where the sample proportion, p, in class C is the usual "unbiased" estimator of the population proportion, P, in class C and $Q = 1 - P$. Thus, for quantitative data p is biased by an amount $B = -P\theta + Q\phi$. This bias is zero only if $\theta = 0$ and $\phi = 0$ or if $-P\theta$ exactly cancels out $Q\phi$.

Now consider the components of variance SV, CC, and SRV in the general variance formula (2.10). It can be shown that

(3.2.2) $$SV = (1 - \lambda)^2\, PQ$$

(3.2.3) $$CC = P^2\sigma^2_\theta + Q^2\sigma^2_\phi - 2PQ\sigma_{\theta\phi}$$

and

(3.2.4) $$SRV = P\theta(1 - \theta) + Q\phi(1 - \phi)$$

where $\lambda = \theta + \phi$, the sum of the average misclassification rates.

These expressions for the components of $V(p)$ lead to entirely different interpretations of the measures of response variance. Consider SV given by (3.2.2). Note that $SV = 0$ if $P = 1$ (or 0)—that is, everyone in the population is (is not) in class C—<u>or</u> if $\lambda = 1$. For example, if both θ and ϕ are $.5$ (and $\lambda = 1$), then we would say the interviewers are equally likely to classify a unit as a "1" or "0" without regard to the unit's true classification. Thus, the proportion of 1's in the population has no effect on sampling variance, and $SV = 0$. Sampling variance is at its maximum when $\lambda = 0$ and $SV = PQ = V(\mu_{ij})$. Thus, misclassification errors reduce sampling variance.

Now consider SRV, the simple response variance. Note that SRV is at minimum when ϕ and θ are both zero (or one), and at a maximum when ϕ and θ are both $.5$. This is just the reverse of the effect of these values on SV. The index of inconsistency

$$I = \frac{SRV}{SV + SRV}$$

is at its maximum (1.0) when $SV = 0$. In most practical situations, this occurs when θ and ϕ are both .5; however, this maximum is also attained whenever $\lambda = 1$. For example, when $\phi = .3$ and $\theta = .7$, or $\phi = .6$ and $\theta = .4$, etc.

The index of inconsistency, therefore, measures the impact of misclassification errors on the total variance of an observation. However, I is not a direct measure of misclassification error. Moreover, it does not account for the effects of correlated errors as in the case of quantitative data as was seen from equations (3.1.8) and (3.2.4). The following example should help in understanding the index I.

Figure a considers hypothetical situations where the errors of misclassification are offsetting or compensating: i.e., units are misclassified with the same frequency both in-category (false positive) and out-of-category (false negatives). Reading across rows of the table, we view the rate of increase in I as the level of random error $\lambda = \theta + \phi$ increases to the maximum; reading down columns, we see that the importance of the error level varies according to the prevalence of the characteristic being measured in the

Figure a. *THEORETICAL VALUES OF I FOR COMPENSATING FALSE POSITIVES AND FALSE NEGATIVES*

($\theta = \phi$)

P	I			
	$\theta=\phi=.01$	$\theta=\phi=.1$	$\theta=\phi=.3$	$\theta=\phi=.4$
.001	.912	.993	.9992	1.000
.01	.510	.934	.9925	.998
.1	.103	.610	.9358	.985
.2	.061	.468	.8913	.974
.3	.047	.401	.8621	.966
.4	.041	.369	.8454	.962
.5	.040	.360	.8400	.960
.6	.041	.369	.8454	.962
.7	.047	.401	.8621	.966
.8	.061	.468	.8913	.974
.9	.103	.610	.9358	.985
.99	.510	.934	.9925	.998
.999	.912	.993	.9992	1.000

population. For example, a combined error rate of 2 percent when $P = .01$ is more serious, as measured by I ($I = .510$), than a combined error rate of 20 percent when $P = .2$ ($I = .468$).

Figure b considers situations where misclassification errors occur primarily in one direction, for example, only false negatives. Because of the direction of the errors, a substantial error rate, say $\theta = .1$, can be tolerated when the units prone to misclassification are not dominant in the population ($P \leq .5$).

These tables indicate the complexities involved in interpreting I. This measure is not necessarily an indicator of the quality of the interview or the interviewers' performance. It measures the reliability of the data and the seriousness of the errors relative to the overall precision of the data.

Figure b. *THEORETICAL VALUES OF I FOR NONCOMPENSATING FALSE POSITIVES AND FALSE NEGATIVES*

($\theta > 0$, $\phi = 0$)

P	I				
	$\theta=.01$ $\phi=0$	$\theta=.1$ $\phi=0$	$\theta=.3$ $\phi=0$	$\theta=.4$ $\phi=0$	$\theta=.5$ $\phi=0$
.001	.010	.100	.300	.400	.500
.01	.010	.101	.302	.402	.503
.1	.011	.101	.323	.426	.526
.2	.013	.122	.349	.455	.556
.3	.014	.137	.380	.488	.588
.4	.017	.156	.417	.526	.625
.5	.020	.182	.462	.571	.667
.6	.025	.217	.517	.625	.714
.7	.033	.270	.588	.690	.769
.8	.048	.357	.682	.769	.833
.9	.092	.526	.811	.870	.909
.99	.503	.917	.977	.985	.990
.999	.910	.991	.998	.999	.999

Now let us consider the interpretation of a large or small CC. The correlated component of interviewer variance is a function of the variability between interviewers in their misclassification probabilities (σ_θ^2 and σ_ϕ^2), the covariance between the probability of a false positive and the probability of a false negative, $\sigma_{\theta\phi}$, and the proportion, P as shown in equation (3.2.3). Assuming that $\sigma_{\theta\phi}$ is nonnegative, which is reasonable for some operations, a large value of CC implies that σ^2_θ and σ^2_ϕ are large, indicating a disparity in the way the operators perform their jobs. A small CC may have several widely differing interpretations. If all the interviewers have the same propensity to cause units in their assignments to be misclassified, CC will be small. This consistency among interviewers is usually regarded as an indication that the operation is under control. This, unfortunately, is not the only condition that will produce a small value for CC. For example, there may be considerable variability between interviewers in the rates at which they misclassify units in class C (large σ^2_θ); however, if the class is rare (small P), the correlated component may still be very small. Therefore, using correlated components as indicators of interviewer performance and quality assurance can be misleading. The following example attempts to exemplify this phenomenon.

Example 3.2--Consider the effect of interviewer misclassification errors on the variance of a categorical survey variable. The impact of correlated nonsampling error on the total variance can be measured by the quantity ρ_y defined above. When the quantitative model is erroneously used for qualitative data, as it often is in practice, a measure ρ^*_y results where $\rho^*_y = \frac{CC}{SV + CC + SRV}$. For this reason, this "hybrid" measure will also be investigated.

Figure c shows ρ_y and ρ^*_y for a series of values of P and for two special cases of the joint distribution of (θ_i, ϕ_i). In the first case, the overall probability of a false negative response is high, say $\theta = .1$, while the overall probability of a false positive response is very small, say $\phi = 0$. Let $\sigma^2_\theta = .005$ which is in the high range for unimodal Beta-distributed, θ_i. In the second case both types of errors occur with high probabilities, $\theta = \phi = .1$, and the population of interviewer misclassification probabilities is highly variable, $\sigma^2_\theta = \sigma^2_\phi = .005$. As a

Figure c. VALUES OF ρ_y AND ρ^*_y FOR GIVEN P, θ, σ^2_θ, ϕ, AND σ^2_ϕ

Population proportion	Case 1: High θ, σ^2_θ; Low ϕ, σ^2_ϕ (θ=.1, σ^2_θ=.005) ($\phi=\sigma^2_\phi$=0)		Case 2: High θ, σ^2_θ; High ϕ, σ^2_ϕ (θ=.1, σ^2_θ=.005) (ϕ=.1, σ^2_ϕ=.005)	
P	ρ_y	ρ^*_y	ρ_y	ρ^*_y
.001	0	0	.0550	.0521
.01	0	0	.0508	.0484
.1	.0006	.0006	.0277	.0270
.2	.0013	.0013	.0176	.0173
.3	.0022	.0022	.0129	.0127
.4	.0034	.0034	.0106	.0105
.5	.0050	.0050	.0100	.0099
.6	.0072	.0071	.0106	.0105
.7	.0105	.0104	.0129	.0127
.8	.0158	.0156	.0176	.0173
.9	.0263	.0256	.0277	.0270
.99	.0504	.0480	.0508	.0484
.999	.0550	.0521	.0550	.0521

further simplification, the covariance is ignored, and the finite population correction is taken to be unity in both examples. For the case where the interviewer probabilities of a false positive (ϕ_i) are high and variable while θ is zero, replace P by $1-P$ in case 1 in figure c. The data in the figure refer to an estimate of P for an area the size of one interviewer's assignment, which is taken arbitrarily as $m = 50$ units.

Figure c dramatically illustrates the effect of the population proportion on the interpretation of the correlated component. Without changing the quality of the interviewer's work--that is, holding the interviewer error parameters constant--very different values of ρ_y can result depending upon the prevalence of the characteristic being measured. In case 1, where only one type of misclassification error is made, ρ_y increases as the proportion of units for which the error is made increases. For case 2, where both types of misclassification errors are possible and are made with identical distributions, ρ_y may still vary considerably according to the magnitude of P.

It is important to realize that, in the case of qualitative data, the correlated component of interviewer variance is not necessarily a measure of the differences among the interviewer systematic errors, as in the case for quantitative data. Therefore, its usefulness as an indicator of data quality is questionnable. However, it is still useful to measure how the total variance of an estimator is affected as a result of systematic interviewer errors.

3.3 *Error structure summary*

The error structures just explored for quantitative and qualitative data as well as the general error formulas are summarized in figure d.

4. THE ESTIMATION OF RESPONSE ERROR COMPONENTS

Two techniques for estimating the components of response error have been widely used in census and survey evaluation: reinterview surveys and interpenetration studies. There are two basic types of reinterview surveys. One type aims at replicating the census or survey interview, using the same procedures, training, questionnaire, and interviewers. This type of reinterview has the objective of estimating *SRV*, the simple response variance. The second type of reinterview survey aims at getting the most correct answer possible and, therefore, may use very expensive procedures, highly qualified and trained interviewers, and detailed, probing questionnaires. This type of reinterview has the objective of estimating B, the response bias. The response bias reinterview is used for estimating B when administrative records are either not available or are infeasible.

Interpenetration studies aim at estimating *CC*, the correlated component of interviewer variance. This procedure randomizes the assignment of households to interviewers so that each interviewer assignment has the same expected value for the average of some characteristic. This procedure is identical to the analysis of variance *(ANOVA)* method of estimating variance components. The next section describes the methodology for estimating *CC*. Reinterview surveys will be addressed in section 4.2.

4.1 *Interpenetration study methodology*

The estimation of *CC* and ρ_y will be discussed for two general designs for interviewer assignment interpenetration. The first design is the <u>full</u> interpenetration of interviewer assignments design. The second design is a two-staged interpenetration design. Both designs are applicable to either quantitative

Figure d. SUMMARY OF NONSAMPLING ERROR STRUCTURES

Symbol	Symbol definition	General form	Quantitative form	Qualitative form
B	Response bias	$\frac{1}{N}\sum_{j=1}^{N} E(e_{j\|j})$	$B_b + B_\varepsilon$	$-P\theta + Q\phi$
SV	Sampling variance	$\underset{j}{V}\, E(e_{j\|j})$	σ^2_μ	$PQ\ (1-\lambda)^2$
CC	Correlated component of enumerator variance	$Cov(e_j, e_{j'}) - Cov(e_a, e_b)$	σ^2_b	$P^2\sigma^2_\theta + Q^2\sigma^2_\phi - 2PQ\sigma_{\theta\phi}$
SRV	Simple response variance	$\frac{1}{N}\sum_{j=1}^{N} V(e_{j\|j})$	$\sigma^2_b + \sigma^2_\iota$	$P\theta(1-\theta) + Q\phi(1-\phi)$
$MSE(Y)$	Mean square error of a census total	$N^2B^2 + N(m-1)\ CC + N\ SRV$	(obtained by appropriate substitution of quantitative components)	(obtained by appropriate substitution of qualitative components)
$MSE(\bar{y})$	Mean square error of a sample mean	$B^2 + (1 - \frac{n-1}{N-1})\frac{SV}{n} + \frac{m-1}{m}\frac{CC}{k} + \frac{SRV}{n}$	↓	↓
I	Index of inconsistency	$\frac{SRV}{SV + SRV}$	↓	↓
ρ_y	Intra-enumerator correlation coefficient	$\frac{CC}{SV + SRV}$	↓	↓

or qualitative data. In each case below, the sample design is assumed to be a simple random sample of units. For more complex survey designs, the estimators of CC are still applicable; however, the estimators of ρ_y are not, since the denominator must be properly weighted in order to estimate $SV + SRV$.

Each estimator described below will, under certain conditions, provide an unbiased estimate of the quantity

$$CC = (Cov(y_{ij}, y_{ij'}) - Cov(y_{ij}, y_{i'j'})$$

This quantity, under quantitative model assumptions, is σ^2_b and under qualitative model assumptions, is $P^2\sigma^2_\theta + Q^2\sigma^2_\phi - 2PQ\sigma_{\theta\phi}$.

4.11. Full interpenetration.--Suppose there are k interviewers or enumerators whose assignments are to be interpenetrated. For full interpenetration, the units in their combined assignments are split as evenly as possible into k random subsamples. Let m_i denote the number of units assigned with simple random sampling to interviewer i where $\sum_1^k m_i = n$, the number of units in all interpenetrated assignments combined. The usual *ANOVA* table for the computations is given in figure e, where

$$(4.1.1) \qquad y_i = \sum_{j=1}^{m_i} y_{ij},\ y = \sum_{i=1}^{k} y_i$$

Figure e. ANALYSIS OF VARIANCE TABLE FOR COMPUTING $\hat{CC}$

Source	Degrees of freedom	Sum of Squares (SS)	Mean Squares (MS)
Among operators (interviewers)	$k - 1$	Operator Sum of Square (OSS) = $\sum_{i=1}^{k} \frac{y_i^2}{m_i} - \frac{y^2}{n}$	Operator Mean Square (OMS) = $\frac{OSS}{k-1}$
Within operators (interviewers)	$n - k$	Error Sum of Squares (ESS) = $\sum_{i=1}^{k} \sum_{j=1}^{m_i} y_{ij}^2 - \sum_{i=1}^{k} \frac{y_i^2}{m_i}$	Error Mean Square (EMS) = $\frac{ESS}{n-k}$

For binomial variables, *OSS* and *ESS* in figure e simplify to

$$OSS = \sum_{i=1}^{k} m_i p_i - np^2$$

$$ESS = npq - OSS$$

where p_i is the proportion in class C in the i^{th} interpenetrated interviewer assignment. Then, an estimator of CC is

$$(4.1.2) \qquad \hat{CC} = \frac{OMS - EMS}{m}$$

where

$$(4.1.3) \qquad m = \sum_{i=1}^{k} m_i^2 \left(\frac{\frac{1}{m_i} - \frac{1}{n}}{k - 1} \right)$$

If the assignment workloads m_i do not vary much between interviewers, instead of (4.1.3), let $m = n/k$.

An estimator of ρ_y can be obtained from the analysis of variance table. For quantitative data the estimator is

$$(4.1.4) \qquad \rho_y = \frac{\hat{CC}}{EMS + \hat{CC}}$$

and for qualitative data the estimator is

$$(4.1.5) \qquad \hat{\rho}_y = \frac{\hat{CC}}{EMS}$$

It can be shown that, if the m_i are treated like constants, the expectation of $\hat{CC}$ is

$$(4.1.6) \qquad E(\hat{CC}) = CC + 2\left[Cov(\mu_{ij}, e_{ij'}) - Cov(\mu_{ij}, e_{i'j'})\right]$$

using the notation of section 3. Thus, $\hat{CC}$ is, in general, a biased estimator of CC. The bias is due to the correlation between the interviewer error and the true values of the characteristics for units within an interviewer's assignment, after accounting for this correlation for units outside of the interviewer's assignment. For example, part of the error e_{ij} may be due to the imputation of non-responses, and this error may be correlated with the true values in another interviewer's assignment. Hence, $Cov(\mu_{ij}, e_{i'j'})$ may not be zero. In addition, the error e_{ij} may be due to an interviewer's personal bias which may be influenced by the values of other units in the same assignment. As mentioned in section 3, such a situation occurred in a survey conducted by the U.S. Census Bureau in the 1950's where the interviewers were to estimate the value of owner-occupied dwellings in the area. The estimated values of the dwellings were influenced by the types of housing units included in the assignment. Hence, the bias term in (4.1.6) may not be zero and must be considered when attempting to interpret the estimated correlated component.

Two examples are given illustrating the estimation of the CC and ρ_y for full interpenetration. In both examples, there are 3 enumerators whose assignments were interpenetrated among 30 units. The sample of 30 units was split as evenly as possible into 3 random subsamples, resulting in 10 units being assigned to each enumerator (i). Data on household income were obtained. In the first example, the income values do not differ much between enumerators, and the estimate of CC and ρ_y is negligible. This indicates no evidence of an intra-interviewer correlation. In the second example, the estimate of CC and ρ_y is small but positive.

Example 4.1--The estimate of CC for the example data provided in figure f is:

$$\hat{CC} = \frac{OMS - EMS}{m} = \frac{51.735 - 216.415}{10} = -16.468$$

And the estimate of ρ_y is:

$$\hat{\rho}_y = \frac{\hat{\sigma}_G^2}{EMS + \hat{\sigma}_0^2} = \frac{-16.468}{216.415 - 16.468} = -.08$$

Negative estimators of CC sometimes occur in practice because the estimator is the difference between two terms. The usual procedure for negative estimates is to replace $\hat{CC}$ and $\hat{\rho}$ by 0 in the reported results.

Example 4.2--The estimate of CC for the example data provided in figure f is:

$$\hat{CC} = \frac{OMS-EMS}{m} = \frac{1,034.8-816.63}{10} = \frac{218.17}{10} = 21.8$$

The estimate of ρ_y is given by:

$$\hat{\rho}_y = \frac{\hat{\sigma}_b^2}{EMS + \hat{\sigma}_b^2} = \frac{21.8}{816.63 + 21.8} = \frac{21.8}{838.43} = .026$$

4.12 Two-staged interpenetration.--A more complex design that is often used for interpenetrating the assignments of interviewers working at a number of different collection offices is the two-staged interpenetration design.

For the first stage, the sample is divided into k mutually exclusive and exhaustive enumeration areas (EA's). These will usually be defined geographically. Let the sample of units within each EA be assigned originally to one and only one interviewer; for example, this might be the area that is most accessible to the interviewer. Now let the k EA's formed for the k interviewers be grouped together arbitrarily into L nonoverlapping "groups" each containing the same number, a, for EA's (see figure f). For example, geographically neighboring EA's may be paired together, $a = 2$, so that the total number of groups is $L = k/2$. Next, a random sample of ℓ groups is drawn from the L groups of EA's.

For the second stage, the interviewers working in these ℓ EA-groups, labelled $h=1, \ldots, \ell$, say, will have their assignments interpenetrated. The interviewers working in the remaining groups, say $h=\ell + 1, \ldots, L$, maintain their original assignments. Thus, if each interviewer assignment is of size m, interviewers in interpenetrated EA's are assigned m/a units at random in each of the a EA's of this group, while interviewers in noninterpenetrated EA's maintained their original assignments of m units in each EA.

Only the EA's in the ℓ groups that are interpenetrated will be used in the estimation. Let y_{hirj} denote the recorded value for the j^{th} unit $(j=1, \ldots, m/a)$ in the (i,r) for $i, r=1, \ldots, a$ interviewer by EA cell for group h of the interpenetrated groups $(h \leq \ell)$. Then, an estimator of CC for interviewers is given by (Biemer and Stokes, 1985).

Figure f. ESTIMATION OF CC

Given data for example 4.1

Interviewers	Household Income (1,000's)										Total
1	20	15	18	12	7	22	30	33	45	60	262
2	25	19	8	17	20	31	16	10	30	42	218
3	61	7	46	28	16	17	18	11	19	27	250

ANOVA table for example 4.1

Source	Degrees of freedom	Sum of Squares	Mean Squares
Among operators (interviewers)	3-1=2	Operator Sum of Square *(OSS)* = 17,866.8 - 17,763.33 = 103.47	Operator Mean Square *(OMS)* = $\frac{103.47}{2}$ = 51.735
Within operators (interviewers)	30-3=27	Error Sum of Squares *(ESS)* = 23,710 - 17,866.8 = 5,843.2	Error Mean Square *(EMS)* = $\frac{5,843.2}{27}$ = 216.415

Given data for example 4.2

Interviewers	Household Income (1,000's)										Total
1	100	32	38	22	45	65	78	10	50	15	455
2	96	22	27	14	38	60	67	3	42	6	375
3	45	50	105	22	60	79	85	28	72	31	577

ANOVA table for example 4.2

Source	Degrees of freedom	Sum of Square	Mean Squares
Among operators (interviewers)	3-1=2	Operator Sum of Squares *(OSS)* = 68,057.9 - 65,988.3 = 2,069.6	Operator Mean Square *(OMS)* = $\frac{2,069.6}{2}$ = 1,034.8
Within operators (interviewers)	30-3=27	Error Sum of Squares *(ESS)* = 90,107 - 68,057.9 = 22,049.1	Error Mean Square *(EMS)* = $\frac{22,049.1}{27}$ = 816.63

(4.1.7) $$\hat{CC} = \frac{1}{m\ell[a(m-2)+1]}\left[\frac{a(m-1)}{(a-1)m} T_I - T_W\right]$$

where

(4.1.8) $$T_I = \sum_{h=1}^{\ell}\sum_{i=1}^{a}(y_{hi..} - m\bar{y}_{h...})^2$$

(4.1.9) $$T_W = \sum_{h=1}^{\ell}\sum_{r=1}^{a}\sum_{i=1}^{a}\sum_{j=1}^{m/a}(y_{hirj} - \bar{y}_{h.r.})^2$$

(4.1.10) $$y_{hi..} = \sum_{r=1}^{a}\sum_{j=1}^{m/a} y_{hirj}$$

(4.1.11) $$m\bar{y}_{h.r.} = \sum_{i=1}^{a}\sum_{j=1}^{m/a} y_{hirj}$$

(4.1.12) $$\bar{y}_{h...} = \frac{\sum_{i=1}^{a} y_{hi..}}{ma}$$

and $n = \ell am$.

An estimator of the denominator of ρ_y for qualitative data given in Biemer and Stokes (1985) is:

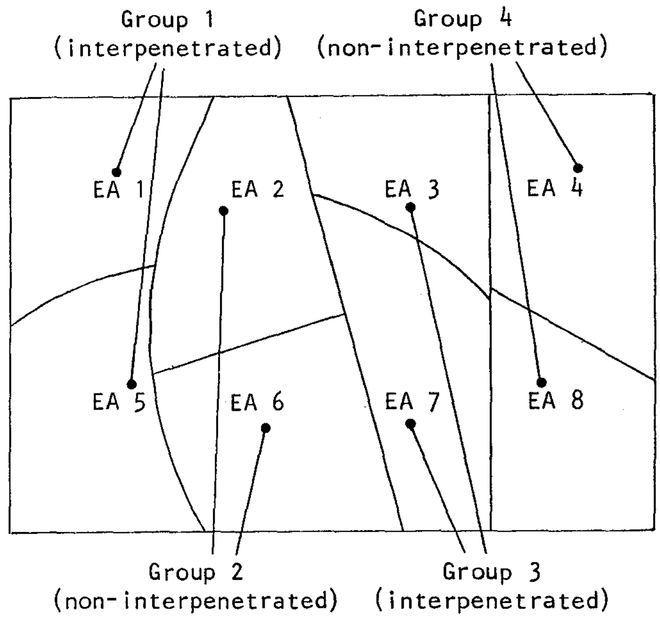
(4.1.13) $$\hat{D} = \frac{1}{\ell[a(m-2)+1]}\left[T_W - \frac{1}{m} T_I\right]$$

$\hat{D}$ has expectation $SV + SRV$ for qualitative data and expectation $SV + SRV - CC$ for quantitative data. Therefore, for quantitative data we use $\hat{D} + \hat{CC}$ to estimate the denominator of ρ_y.

Example 4.3--Consider a two-staged interpenetration as illustrated in figure g for 4 interviewers (A,B,C, and D). Interviewers A and B had their assignments paired: EA 1 and EA 5, respectively, are interpenetrated. Likewise, interviewers C and D had their assignments interpenetrated: EA's 3 and 7, respectively. All remaining assignments were not interpenetrated, and their results will not enter into the subsequent computations.

Each interviewer was assigned 50 housing units and was instructed to classify each housing unit as either above or below the poverty level. Interviewer A classified 20 percent of the housing units in his/her

Figure g. TWO-STAGED INTERPENETRATION DESIGN FOR $k = 8$, $a = 2$, AND $\ell = 2$



interpenetrated assignment below the poverty level. Interviewer B classified 60 percent below the poverty level. Interviewer C classified 30 percent and interviewer D classified 10 percent of their assigned housing units below the poverty level. The following represents the complete results:

$y_{11..} = 50(.2) = 10$; the number of households interviewer A classified below poverty level.

$y_{12..} = 50(.6) = 30$; the number of households interviewer B classified below poverty level.

$y_{21..} = 50(.3) = 15$; the number of households interviewer C classified below poverty level.

$y_{22..} = 50(.1) = 5$; the number of households interviewer D classified below poverty level.

$m = 50$; the interviewer assignment size.

$\bar{y}_{1...} = \frac{10 + 30}{100}$ = the proportion of households classified below poverty level in Group 1.

$\bar{y}_{2...} = \frac{15 + 5}{100}$ = the proportion of households classified below poverty level in Group 2.

$$T_I = \sum_{h=1}^{2} \sum_{i=1}^{2} (y_{hi..} - m\bar{y}_{h...})^2$$

$$= (10-20)^2 + (30-20)^2 + (15-10)^2 + (5-10)^2$$

$$= 250$$

$y_{111.}$ = 5 = the number of households classified below poverty level by interviewer A in the first EA (EA1) of the first group.

$y_{112.}$ = 5 = the number of households classified below poverty level by interviewer A in the second EA (EA5) of the first group.

$y_{211.}$ = 7 = the number of households classified below poverty level by interviewer C in the first EA (EA3) of the third group.

$y_{212.}$ = 8 = the number of households classified below poverty level by interviewer C in the second EA (EA7) of the third group.

$y_{121.}$ = 15 = the number of households classified below poverty level by interviewer B in the first EA (EA1) of the first group.

$y_{122.}$ = 15 = the number of households classified below poverty level by interviewer B in the second EA (EA5) of the first group.

$y_{221.}$ = 3 = the number of households classified below poverty level by interviewer D in the first EA (EA3) of the third group.

$y_{222.}$ = 2 = the number of households classified below poverty level by interviewer D in the second EA (EA7) of the third group.

$\bar{y}_{1.1}$ = $\frac{20}{50}$ = the proportion of households classified below poverty level in EA1.

$\bar{y}_{1.2.}$ = $\frac{20}{50}$ = the proportion of households classified below poverty level in EA5.

$\bar{y}_{2.1.}$ = $\frac{10}{50}$ = the proportion of households classified below poverty level in EA3.

$\bar{y}_{2.2.}$ = $\frac{10}{50}$ = the proportion of households classified below poverty level in EA7.

$$T_W = \sum_{h=1}^{2} \sum_{r=1}^{2} \sum_{i=1}^{2} \sum_{j=1}^{m/a} (y_{hirj} - \bar{y}_{h.r.})^2$$

$$= 20(1 - \frac{20}{50})^2 + 30(- \frac{20}{50})^2 + 20(1 - \frac{20}{50})^2 + 30(- \frac{20}{50})^2 + 10(1 - \frac{20}{50})^2 + 40(- \frac{20}{50})^2 + 10(1 - \frac{20}{50})^2 + 40(\frac{20}{50})^2$$

$$= 40$$

An estimator of CC for interviewers is thus,

$$CC = \frac{1}{m\ell[a(m-2)+1]} \left[\frac{a(m-1)}{(a-1)m} T_I - T_W \right]$$

$$= \frac{1}{50(.2)[2(48)+1]} \left[\frac{2(49)}{1(50)} 250 - 40 \right]$$

$$= .0464$$

An estimator of the denominator of ρ_y is

$$\hat{\sigma}_D^2 = \frac{1}{\ell [a(m-2)+1]} \left[T_W - \frac{1}{m} T_I \right]$$

$$= \frac{1}{2[2(48)+1]} \left[40 - \frac{250}{50} \right]$$

$$= .1804$$

so that

$$\hat{\rho}_y = \frac{.0464}{.1804} = .2572$$

This is an extremely large value of ρ_y and would usually indicate that substantial systematic errors were committed by the interviewers. In this case, however, since only four interviewer assignments were used in the estimation, $\hat{\rho}_y$ is subject to considerable sampling variability.

4.13 Precision.--The variance of the estimator of CC for two-staged interpenetration has been derived assuming the quantitative data model of section 3.1 and normally distributed observations, y. Let $\sigma_{\mu'}^2$, denote the average within EA variance of the terms $\mu'_{hirj} = \mu_{hirj} + \varepsilon_{hirj}$ in model (3.1.3).

Then ρ_y has the form

$$(4.1.14) \qquad \rho_y = \frac{\sigma_b^2}{\sigma_{\mu'}^2 + \sigma_b^2}$$

The relative variance *(Relvar)*, which is the coefficient of variation squared, of the estimator (4.1.7) of interviewer variance has been derived and can be expressed in terms of the magnitude of ρ_y. This expression is given approximately by (Biemer and Stokes 1985)

$$(4.1.15) \qquad Relvar\ (\hat{CC}) = \frac{2}{\ell(a-1)}\left[1 + \frac{2}{m}\left(\frac{1-\rho_y}{\rho_y}\right) + \frac{a(m-1)}{m^2(am-2a+1)}\left(\frac{1-\rho_y}{\rho_y}\right)^2\right]$$

It should be recognized that this formula should only be used as a very rough guide. It can be useful for determining the number of interviewer assignments to interpenetrate for some specified level of precision in the estimation $\hat{CC}$.

Example 4.4--Consider the two-staged interpenetration design and suppose $\ell = 259$ groups of $a = 4$ interviewer assignments are chosen for interpenetration. The average assignment size for each interviewer is 125 households. (This design was used in the 1970 U.S. Census Enumerator Variance Study.) Suppose that from previous studies, it is known that, for the characteristics of interest, ρ_y is approximately *.01*. Then the expected relative variance of the estimator of *CC* can be computed using (4.1.15). Under the assumptions stated above, the estimated *Relvar* $(\hat{CC})$ is approximately *.008* or a coefficient of variation of *.09*.

4.2 *Reinterview surveys*

Reinterview surveys can provide estimates of either the simple response variance, *SRV*, or response bias, *B*, or both *SRV* and *B*, depending upon the objectives and the design. The designs we will consider will not allow the separate estimation of *CC*. The simplest way to describe the theory of estimation of *SRV* and *B* is through our general model

$$(4.2.1) \qquad y_j = \mu_j + e_j$$

defined in section 2. This model refers to a single observation of some characteristic c for the unit j. In order to distinguish between the original interview and the reinterview, we will incorporate a double subscript (j,t) to denote the t^{th} observation on unit j. Thus, $t = 1$ denotes the census (or original survey) response, and $t = 2$ denotes the reinterview survey response. We now have two models corresponding to the two observations:

$$(4.2.2) \qquad y_{j1} = \mu_j + e_{j1} \ ,$$

the model for the original interview and

$$(4.2.3) \qquad y_{j2} = \mu_j + e_{j2} \ ,$$

the model for the reinterview. Note that in both models, μ_j is not indexed by t. This reflects the following assumption: the true value, μ_j, for the j^{th} unit in the sample is assumed to be the same for the original interview and the reinterview. This assumption can usually be satisfied for most characteristics. For example, suppose the characteristic, c, is age. We are not saying that the respondents do not age between the census and the reinterview. Rather, we are saying that, in the reinterview, the respondent's age at the time of the census (not the current age) is to be obtained. Therefore, the reinterview questions must be quite different, for some characteristics, to take into account the time lag between the census and the reinterview. Stated simply, the census and the reinterview are intended to measure the same thing--an individual's age, marital status, nationality, etc., for the census reference period.

The errors e_{j_1} and e_{j_2} may, of course, be different for the two interviews. This leads to the following assumption:

$$E(e_{j_1}) = B \text{ and } V(e_{j_1}) = \sigma_e^2$$

$$E(e_{j_2}) = B' \text{ and } V(e_{j_2}) = \sigma_{e'}^2,$$

where B and B' and σ_e^2 and $\sigma_{e'}^2$ are not necessarily equal.

The means and variances of the errors e_{j_1} and e_{j_2} depend upon the general conditions under which the census and the reinterview are taken. Some of these general conditions may be beyond the control or specification of the census and study designers as, for example, the general political, economic, and social situation at the time of the interview, rumors, or general uncontrolled publicity. Uncontrolled conditions also include many temporary chance situations appearing at the time a response is obtained. Some conditions can be controlled to influence the quality of results. These controls, which attempt to ensure adequate quality, affect questionnaire design, interview procedures, certain aspects of census publicity, personnel qualifications, pay, training and inspection, and other controls.

Let G denote the set of general conditions operating during the census, which gives rise to the error distribution for e_{j_1}. Let G' denote the set of conditions operating during the reinterview, which gives rise to the error distribution for e_{j_2}. If $G = G'$, that is, if all general conditions which affect response errors are the same for both the census and the reinterview, then by definition, $B = B'$ and $\sigma_e^2 = \sigma_{e'}^2$. Likewise, if $G \neq G'$, that is, the census and the reinterview survey are conducted under quite different general conditions, then we cannot say that $B = B'$ or $\sigma_e^2 = \sigma_{e'}^2$.

Finally, we assume that the errors e_{j_1} and e_{j_2} are independent; i.e.,

$$Cov(e_{j_1}, e_{j_2}) = 0 \text{ for all } j.$$

In practice, this assumption may not hold as a result of respondent conditioning. Technically, a conditioning effect is present when uncontrollable influences are brought to bear on the respondent (or interviewer) as a direct result of the original interview. For example, the respondent may simply recall previously given responses to questions in the original interview and repeat them in the reinterview. This is affected, to some extent, by the timing of the reinterview. A greater time lag between the census and the reinterview can reduce this effect. This assumption will be examined in more detail later, and the impact of between trial correlations will be discussed. Next, we consider the estimation of B and SRV. First, however, some additional notation is needed. In the following, we assume that the reinterview sample constitutes a simple random sample without replacement of size n from the census (or survey).

4.21 Notation for quantitative data.--For quantitative data, both e_{j_1} and e_{j_2} have the structure given in (3.1.1). Define $\bar{y}_1$ and $\bar{y}_2$ to be the sample means of the census observations and the reinterview observations, respectively. Let

$$2\sigma_R^2 = \frac{\sum_{j=1}^{n} (y_{j_1} - y_{j_2})^2}{n}$$

denote the average sum of squares of differences between the two observations on the same units. Then, in general,

$$E(\bar{y}_1 - \bar{y}_2) = B - B' \tag{4.2.4}$$

and

$$E(\hat{\sigma}_R^2) = \frac{1}{2}(\sigma_b^2 + \sigma_\varepsilon^2 + \sigma_{b'}^2 + \sigma_{\varepsilon'}^2) \tag{4.2.5}$$

where the primes denote the distributions for the reinterview error variables b'_i and ε'_{ij} (using the notation of section 3.1).

4.22 Notation for qualitative data.--From the data collected in the interview and reinterview, the following table shown as figure h can be generated.

Figure h. CROSS-CLASSIFICATION OF ORIGINAL INTERVIEW RESULTS BY REINTERVIEW RESULTS

Reinterview (y_{j2})	Original interview (y_{j1}) 1	0	Total
1	a	b	a + b
0	c	d	c + d
Total	a + c	b + d	n

For example, a is the number of units classified in C for both the original survey and the reinterview, b is the number classified as not in C in the survey but in C in the reinterview, etc. The underlying population parameters of these data are represented by a similar table shown in figure i.

Figure i. PROBABILITIES FOR THE ORIGINAL INTERVIEW AND REINTERVIEW CROSS-CLASSIFICATION

Reinterview (y_{j2})	Original interview (y_{j1}) 1	0	Total
1	π_a	π_b	π'
0	π_c	π_d	$1 - \pi$
Total	π	$1 - \pi$	1

That is, the expected values of the cell totals in figure h are n times the cell parameters in figure i.

The symbols in the table may be interpreted as follows: Suppose a unit is drawn at random from the population; π is the probability that the unit is classified in C in the original survey, while π' is the probability that the unit is classified in C in the reinterview survey. The probability that the unit is classified in C on both occasions is π_a and classified not in C on both occasions is π_d. The probability that a unit is classified as not belonging to C in the original survey and in C in the reinterview is π_b, while the probability that the original survey classifies the unit in C but the reinterview classifies the unit not in C as π_c. It may be shown that cells in figure h, divided by n unbiasedly, estimate the corresponding cells in figure i. This can be expressed mathematically as

$$(4.2.6)\quad \pi_a = E(\frac{a}{n}) = P\ (1 - \theta)\ (1 - \theta') + Q\ \phi\phi'$$

$$(4.2.7)\quad \pi_b = E(\frac{b}{n}) = P\ (1 - \theta')\theta + Q\ (1 - \phi)\phi'$$

$$(4.2.8)\quad \pi_c = E(\frac{c}{n}) = P\ (1 - \theta)\theta' + Q\ \phi\ (1 - \phi')$$

$$(4.2.9)\quad \pi_d = E(\frac{d}{n}) = P\ \theta\theta' + Q\ (1 - \phi)\ (1 - \phi')$$

$$(4.2.10)\quad \pi = E(\frac{a+c}{n}) = P\ \theta'\ (1 - \theta) + Q\ \phi\ (1 - \phi')$$

$$(4.2.11)\quad \pi' = E(\frac{a+b}{n}) = P\ \theta\ (1 - \theta') + Q\ \phi'\ (1 - \phi)$$

Define the net difference rate as the difference between the estimates of P from the two trials, say p_1 and p_2, given by $p_1 - p_2$. The expected value of the net difference rate, in general, is

$$(4.2.12)\quad E(p_1 - p_2) = E\left(\frac{b-c}{n}\right) = \pi - \pi'$$

$$= -P(\theta - \theta') + Q(\phi - \phi')$$

Define the <u>gross difference rate</u> as the rate of gross disagreement between the two trials denoted by g where

$$(4.2.13)\qquad g = \frac{b+c}{n}$$

and, in general,

$$(4.2.14)\qquad E(g) = \pi_b + \pi_c$$

4.23 <u>Estimation for both types of data</u>.--This section will consider the estimation of *SRV*, *I*, and the bias, *B*, for the census or survey.

Case 1: The Identical Second Trial

(1) <u>Simple response variance for quantitative data</u>. If the conditions at each trial are the same, that is, $G = G'$, then the error components σ_b^2 and σ_ε^2 are the same for both trials and

$$E(\sigma_R^2) = \sigma_b^2 + \sigma_\varepsilon^2 = SRV$$

Let $s_t^2 = \dfrac{\sum_{j=1}^{n} (y_{jt} - \bar{y}_t)^2}{n-1}$

denote the sample variances for the two trials. Then, it can be shown that, for $t = 1,2$

$$E(s_t^2) = \frac{N}{N-1}\,\sigma_\mu^2 + \frac{n-m}{n-1}\,\sigma_b^2 + \sigma_\varepsilon^2$$

An estimator of the index of inconsistency is

$$\hat{I} = \frac{2\hat{\sigma}_R^2}{s_1^2 + s_2^2}$$

In the presence of correlated errors (i.e., $\sigma_b^2 > 0$), $\hat{I}$ may slightly overestimate I.

(2) <u>Simple response variance for qualitative data</u>. If the conditions at each survey trial are identical $(G = G')$ then $\theta = \theta'$ and $\phi = \phi'$, i.e., the probability of misclassification is the same for both surveys. Therefore, $\pi_b = \pi_c = SRV$ from (4.2.6) and (4.2.6) and (4.2.8) and, hence, an unbiased estimator of *SRV* is $g/2$ (from 4.2.13).

The denominator of the index of inconsistency, *SV* + *SRV*, can be unbiasedly estimated by either p_1q_1 or p_2q_2. Thus a consistent estimator of *I* is

$$\hat{I}\quad \frac{g}{p_1q_1 + p_2q_2}$$

This is true even in the presence of correlated errors. Other estimators of *I* can be formed by using $2p_1q_1$ or $p_1q_2 + p_2q_1$ in the denominator.

(3) <u>Bias</u>. No estimator of *B* is possible for either quantitative or qualitative data.

Case 2: The "Improved" Second Trial

Many times $G' \neq G$ because the second trial implements an improved procedure. This is usually the case when our objective is to obtain the "truth" in the reinterview: the best possible result is not the "truth" but a better response than in the original interview.

(1) <u>Simple response variance for quantitative data</u>. Specifically, it is assumed that $\sigma_{b'}^2 < \sigma_b^2$ and $\sigma_{\varepsilon'}^2 < \sigma_\varepsilon^2$ where $\sigma_{b'}^2$ and $\sigma_{\varepsilon'}^2$ are the variance components for the improved procedure. Defining expectation now over all possible trials under <u>both</u> conditions G and G', we have

$$E(\hat{\sigma}_R^2) = \frac{1}{2}(\sigma_b^2 + \sigma_{b'}^2) + \frac{1}{2}(\sigma_\varepsilon^2 + \sigma_{\varepsilon'}^2)$$

$$< \sigma_b^2 + \sigma_\varepsilon^2$$

Thus, $\hat{\sigma}_R^2$ <u>underestimates</u> *SRV*.

Similarily, it can also be shown that

$$\hat{I} = \frac{2\hat{\sigma}_R^2}{s_1^2 + s_2^2}$$

<u>underestimates</u> *I*.

(2) Simple response variance for qualitative data. If the second trial is an improved procedure relative to the original survey, then this means precisely that

$$\theta' < \theta$$

and

$$\phi' < \phi$$

Just as for quantitative data, no unbiased estimator of any of the measures of interest exists. By elementary algebra, using (4.2.7), (4.2.8), and (4.2.13),

$$E(g) < 2\ SRV,$$

that is, the gross difference rate will underestimate the simple response variance for the survey. Further, a lower bound estimator of the index of inconsistency, I, is

$$\hat{I} = \frac{g}{2p_1q_1}$$

(3) Bias for quantitative data. Let B' denote $E(e_{j2})$, the bias in the second trial. Then

$$E(\bar{y}_1 - \bar{y}_2) = B - B'$$

and the difference $\bar{y}_1 - \bar{y}_2$ estimates the differential bias of the two survey procedures. If it may be assumed that $B' \doteq 0$, then $\bar{y}_1 - \bar{y}_2$ provides an approximation to B.

(4) Bias for qualitative data. From (4.2.12), the net difference rate provides an approximation to the bias, B, if $\phi' \doteq \theta' \doteq 0$. Otherwise, the net difference rate estimates the differential bias of the census and is a lower bound estimator for B.

Case 3: The Perfect Second Trial

(1) Simple response variance for quantitative data. Suppose data are available, either through reinterview, administrative records, or otherwise, such that the second measurement represents "truth." Then $\sigma^2_{b'} = \sigma^2_{\varepsilon'} = 0$; that is, there is no nonsampling error for the second measurement and

$$E(\hat{\sigma}^2_R) = \frac{1}{2}\ SRV$$

A consistent estimator of I is

$$\hat{I} = \frac{\hat{\sigma}^2_R}{2s_1^2}$$

(2) Simple response variance for qualitative data. If it may be assumed that the second trial is performed without error, an alternative interpretation of the symbols in figure 1 may be given as follows: For a unit chosen at random from the population, π is the probability that the unit is classified in C regardless of what class it belongs to, while π' is the probability that it truly belongs to C. The probability that the unit belongs in C and is correctly classified is π_a and is incorrectly classified is π_b. Likewise, the probability that a unit does not belong to C and is correctly classified is π_d and incorrectly classified is π_c. Mathematically, this means that $\theta' = 0$, $\phi' = 0$, and therefore, from (4.2.6) to (4.2.11), we now have

$$\pi_a = P(1 - \theta)$$

$$\pi_b = P\ \theta$$

$$\pi_c = Q\ \phi$$

$$\pi_d = Q(1 - \phi)$$

$$\pi = P(1 - \theta) + Q\ \phi$$

$$\pi' = P$$

From these relationships, the parameters, P, θ, and ϕ are all estimable and thus, estimates of SRV and I can be computed by the "method of moments" (i.e., equating the parameter functions with their unbiased estimators and solving for the parameters). For example, an estimator of I is

$$\hat{I} = \frac{\hat{\pi}_a\hat{\pi}_b/\hat{\pi}' + \hat{\pi}_c\hat{\pi}_d/(1 - \hat{\pi}')}{\hat{\pi}(1 - \hat{\pi})}$$

$$= \frac{abc + abd + acd + bcd}{n^3\ p_1q_1p_2q_2}$$

(3) Bias for quantitative data. An unbiased estimator of B is given by $\hat{B} = \bar{y}_1 - \bar{y}_2$.

(4) Bias for qualitative data. An unbiased estimate of B is given by $\hat{B} = p_1 - p_2$, the net difference rate.

4.24 Covariance between trials.--Until now, it has been assumed that the errors e_{j1} and e_{j2} in the general model are independent, or that the surveys were conducted completely independently. Of course, for reinterview surveys, this is merely conceptual. In practice, a survey usually cannot be repeated independently, because respondents have been exposed to the original survey and may merely repeat answers given in the original survey, or their attitudes, perceptions, knowledge, etc., may have changed as a result of their first exposure to the survey. There are techniques for minimizing these correlations; they were discussed in section 2.24 of chapter 3. We now consider the effect of correlation between trials on the estimator of *SRV*.

Consider Case 1 (identical second trial) again and denote the correlation between trials by ρ_T. It is assumed that ρ is the same for all units, then

$$\rho_T = \frac{\underset{j}{E}\, Cov(e_{j1}, e_{j2}|j)}{\underset{j}{E}V(e_{jt}|j)}$$

It can be easily shown that the expected value of $\hat{\sigma}^2_R$ (quantitative data) or $g/2$ (qualitative data) is

$$SRV\ (1 - \rho_T)$$

Therefore, $\hat{\sigma}^2_R$ and $g/2$ will be poor estimators of *SRV* when there is a large correlation among the errors in a census and a reinterview survey. In fact, the usual estimators of *SRV* and I will typically underestimate these parameters, since ρ_T is usually non-negative. Techniques exist for estimating ρ_T; however, these require much more complex evaluation study designs and are not in our scope of work. See (Bailar 1968) and (Hansen et al. 1964) for further discussion.

4.25 Precision of the estimators.--We shall now consider the precision of the estimator

$$\hat{I} = \frac{g}{2p_1q_1}$$

Our study is confined to the qualitative data case, since this case is more commonly encountered in practice.

Since the estimator of the index of inconsistency is a ratio of random variables, no exact and tractable expression of the variance of $\hat{I}$ exists. However, an approximate formula for the variance can be derived which is valid for large reinterview samples. We first give the formulas for the rel-variance (coefficient of variation squared) of the numerator, the denominator, and the relative covariance of the numerator and denominator of $\hat{I}$.

Assume simple random sampling of the population and that the sampling fraction n/N is negligible. Consider the simplest estimator of I, viz.

$$\hat{I} = \frac{g}{2p_1q_1} \tag{4.2.15}$$

It can be shown that

$$Relvar(g) = \frac{\pi_a + \pi_b}{n(\pi_b + \pi_c)} \tag{4.2.16}$$

$$Relvar(2p_1q_1) = \frac{1}{n}\left[\frac{1}{\pi(1-\pi)} - \frac{4n-6}{n-1}\right] \tag{4.2.17}$$

and

$$Relcov(g, 2p_1q_1) = \frac{1}{n}\left[\frac{\pi_c}{\pi_b + \pi_c}\cdot\left(\frac{\pi}{1-\pi}\right) + \frac{\pi_b}{\pi_b + \pi_c}\left(\frac{1-\pi}{\pi}\right) - 1\right] \tag{4.2.18}$$

Therefore, the relative variance of the ratio $\hat{I}$ is approximately:

$$(4.1.19)\quad Relvar\ (\hat{I}) \doteq \frac{1}{n}\ Relvar(g) + Relvar(2p_1q_1) - 2\ Relvar(g, 2p_1q_1)$$

Equations (4.2.15) through (4.2.19) are true regardless of the value of ρ_T or the conditions G and G'. A consistent estimator of (4.2.19) can be obtained by simply replacing the parameter values by their sample estimators. This estimator is

$$(4.2.20)\quad Relvar(\hat{I}) = \frac{1}{n}\left[\frac{1-g}{g} + \frac{1}{p_1q_1} - 2\,\frac{nc}{g}\,\frac{p_1}{q_1} + \frac{nb}{g}\cdot\frac{q_1}{p_1} - \frac{2n-4}{n-1}\right]$$

Estimator (4.2.20) is not used in chapter 3 (section 3), since it is computationally difficult. In figure 3-14 a different estimator of the relative variance of $\hat{I}$ is suggested, that is, a combination of two alternative estimators. When g is $.10$ or less, an estimator is used which is based upon the assumptions that (a) g is distributed as a Poisson random variable, and (b) the denominator of $\hat{I}$ is constant. The form of the estimator is

$$(4.2.21)\quad Relvar(\hat{I}) = \frac{1}{ng}$$

When g is greater than $.10$, an estimator is used that replaces assumption (a) above by *(a')*, resulting in g having the binomial distribution. The estimator in this case is

$$(4.2.22)\quad Relvar(I)\quad \frac{(1-g)}{ng}$$

It should be noted that figure 3-14 includes adjustments ($g+2$ in place of g and $\frac{1}{n(g+1)}$ in place of $\frac{1}{ng}$) to improve the properties of the estimated confidence intervals.

Example 4.5--To illustrate the use of these formulas, consider the data in figure j on vocational training from the 1970 U.S. Census Reinterview for estimating *SRV*. For this table, $n = 7{,}567$, $a = 859$, $b = 556$, $c = 524$, $d = 5{,}628$, and $p_1 = \frac{1{,}383}{7{,}567} = .18$. Thus, an estimate of I using (4.2.15) is

$$\hat{I} = \frac{556 + 524}{2(7{,}567)} \cdot \frac{1}{(.18)(.82)} = .48$$

and, using (4.2.20), an estimate of the relative variance of $\hat{I}$ is

$$Relvar(\hat{I}) = \frac{1}{7{,}567}\left[5.21 + 6.78 - (.32)\right]$$

$$= .00167$$

or an estimated coefficient of variation of about $\sqrt{.00167}$ or 4 percent which indicates good precision of the estimate of I. Using the computationally efficient estimator of the $Relvar(\hat{I})$, (4.2.22) applies, since $g = .14$ which is greater than $.1$. Thus

$$Relvar(\hat{I}) = \frac{(1 - .14)}{(7{,}567).14}$$

$$= .000817$$

or on estimated $CV(\hat{I})$ of 3 percent.

Figure j. *DISTRIBUTION OF CENSUS AND REINTERVIEW RESPONSES FOR VOCATIONAL TRAINING*

(Data shown as numbers of sample cases)

Reinterview classification	Census response		
	Total sample persons	Completed a vocational training program	Did not complete a vocational training program
Total sample persons......	7,567	1,383	6,184
Completed a vocational training program........	1,415	859	556
Did not complete a vocational training program.......	6,152	524	5,628

5. REFERENCES

Bailar, B.A. 1968. "Recent Research in Reinterview Procedures." *Journal of the American Statistical Association* 63:41-63.

Bailar, B.A., and Biemer, P.P. 1984. Some Methods for Evaluating Nonsampling Error in Household Censuses and Surveys." In *W.G. Cochran's Impact on Statistics*, edited by P.S.R.S. Roo, 253-274. New York: John Wiley and Sons, Inc.

Biemer, P.P., and Stokes, S.L. 1985. "Optimal Design of Interviewer Variance Experiments in Complex Surveys." *Journal of the American Statistical Association* 80:158-166.

Hansen, M.H.; Hurwitz, W.N.; and Pritzker, L. 1964. "The Estimation and Interpretation of Gross Differences and the Simple Response Variance." In *Contributions to Statistics*, edited by C.R. Rao, 111-136. Oxford, Pergamon Press.

GLOSSARY

A

ACCURACY: Range of census variation or the degree of conformity of the census count to the actual number of persons or housing units.

AGE-HEAPING: Tendency for enumerators or respondents to report certain ages instead of others; also known as age preference or "digit preference." Preference for ages ending in zero or five is widespread.

AGE PATTERN OF FERTILITY: Relative distribution of a set of "age-specific fertility rates." It expresses the relative contribution of each age group to "total fertility."

AGE RATIO: Ratio of the population in a given age group to the average of the populations in the two neighboring age groups, times 100.

AGE-SPECIFIC FERTILITY RATE: Number of births occurring during a specified period to women of a specified age or age group, divided by the number of person-years lived during that period by women of that age or age group. When an age-specific fertility rate is calculated for a calendar year, the number of births to women of the specified age is usually divided by the mid-year population of women that age.

AGE-SPECIFIC MORTALITY RATE: Number of deaths occurring during a specified period to persons (usually specified by sex) of a specified age or age group, divided by the number of person-years lived during that period by the persons of that age or age group. When an age-specific mortality rate is calculated for a calendar year, the number of deaths to persons of the specified age is usually divided by the mid-year population of persons of that age. Age-specific mortality rates are generally denoted by ${}_nM_x$, the annual death rate of persons aged from x to $x + n$.

AGGREGATES: Individual observations (or statistics) combined into a total number.

ASSIGNABLE VARIATION: Variation attributable to causes other than chance.

B

BALANCING EQUATION: One technique for evaluating a census. The expected population at the present census equals the population at the last census plus the births and immigrants during the intercensal period minus the deaths and emigrants during the intercensal period.

BIAS: Difference between the expected value of a statistic and its true value.

C

CHANDRASEKARAN-DEMING TECHNIQUE: Procedure to estimate the coverage of two independent systems collecting information about demographic or other events, based on the assumption that the probability of an event being recorded by one system is the same regardless of whether the event is recorded by the other system. The events from both systems are matched to establish M, the number of events recorded by both systems; U_1, the number recorded only by system 1; and U_2, the number recorded only by system 2. The Chandrasekaran-Deming formula then estimates total events, N, as

$$\hat{N} = M + U_1 + U_2 + U_1 U_2 / M.$$

CHILDBEARING AGES: The span within which women are capable of bearing children, generally taken to be from age 15 to age 49 or, sometimes, to age 44.

CHILDREN EVER BORN: Number of children ever born alive by a particular woman; synonymous with "parity." In demographic usage, stillbirths are specifically excluded.

COEFFICIENT OF VARIATION (CV): Standard error divided by the population parameter being estimated; another name for "relative standard error."

COHORT: Group of persons who experienced the same class of events in the same period. Thus, an age cohort is a group of people born during a particular period, and a marriage cohort is a group of people who married during a particular period. The effects of a given set of mortality or fertility rates are often illustrated by applying them to hypothetical cohorts.

COHORT FERTILITY: The fertility experienced over time by a group of women or men who form a birth or marriage cohort.

CONDITIONING BIAS: Tendency in a subsequent interview to repeat the original response for a question whether or not it is perceived as the correct answer.

CONFIDENCE INTERVAL: Interval within one standard error (two-thirds confidence interval) of the estimate or two standard errors (95 percent confidence interval).

CONTENT ERROR: Error in the characteristics that are reported for those persons or housing units that are enumerated.

CORRELATION BIAS: Result of failure to achieve independence between the complete enumeration and the coverage evaluation survey.

COVERAGE ERROR: Error in an estimate that results from (a) failure to include in a census or a survey all eligible units or (b) inclusion of some units erroneously.

COVERAGE EVALUATION: Process by which data are collected for a sample of households after the census for the purpose of evaluating completeness of the census enumeration.

CRUDE BIRTH RATE: Number of births in a population during a specified period divided by the number of person-years lived by the population during the same period. It is frequently expressed as births per 1,000 population. The crude birth rate for a single year is usually calculated as the number of births during the year divided by the mid-year population.

CRUDE DEATH RATE: Number of deaths in a population during a specified period divided by the number of person-years lived by the population during the same period. It is frequently expressed as deaths per 1,000 population. The crude death rate for a single year is usually calculated as the number of deaths during the year divided by the mid-year population.

CUMULATED FERTILITY: An estimate of the average number of children ever born by women of some age x, obtained by cumulating "age-specific fertility rates" up to age x; also often calculated for age groups.

D

DE FACTO POPULATION: Population enumerated on the basis of those present at a particular time, including temporary visitors and excluding residents temporarily absent.

DE JURE POPULATION: Population enumerated on the basis of normal or usual residence, excluding temporary visitors and including residents temporarily absent.

DUAL SYSTEM ESTIMATION: Preparation of estimates by matching, on a case-by-case basis, two different and independent sources which describe the same events. The matching permits an estimate of the number of cases reported by one source but not the other.

E

ENUMERATOR: Person who collects data.

EXPECTATION OF LIFE AT BIRTH: Average number of years that a member of a "cohort" of births would be expected to live if the cohort were subject to the mortality conditions expressed by a particular set of "age-specific mortality rates." Denoted by the symbol e_0 in "life table" notation.

EXPECTED VALUE: Average value of sample estimates over all independent repetitions.

F

FORWARD SURVIVAL: Procedure for estimating the age distribution at some later date by projecting forward an observed age distribution. The procedure uses "survival ratios," often obtained from model "life tables." The procedure is basically a form of population projection without the introduction of new entrants (births) to the population.

FREQUENCY DISTRIBUTION: Distribution of estimates from a sample.

G

GROSS ERROR: Census count for any category before subtracting the correct count for that category.

GROWTH RATE: The increase or decrease of a population in a period divided by the number of person-years lived by the population during the same period. The increase in a population is the result of a surplus (or deficit) of births over deaths and of immigrants over emigrants.

H

HOUSEHOLD: A person or group of related or unrelated persons who make common provision for food and other necessities for living.

HOUSING UNIT: A single room or group of rooms, or other space arranged for human habitation, occupied or intended for occupancy as separate and independent living quarters by a person living alone or a group of persons living together.

I

INFANT MORTALITY RATE: Number of deaths of children under one year of age occurring in the same year; also used in a more rigorous sense to mean the number of deaths that would occur under one year of age in a "life table" with a "radix" of 1,000, in which sense it is denoted by the symbol ${}_1q_0$.

IMPUTATION: Process whereby missing or inconsistent data are replaced by estimated values that are derived by considering other related data items that are present.

L

LIFE TABLE: Listing of the number of survivors at different ages (up to the highest age attained) in a hypothetical "cohort" subject from birth to a particular set of "age-specific mortality rates." The rates are usually those observed in a given population during a particular period of time. The survivors of the "radix" to age x are generally denoted by $\ell(x)$. The tabulations commonly accompanying a life table include other features of the cohort's experience: its expectation of life at each age x, denoted by e_x; the probability of dying from each age x to age $x + n$, denoted by ${}_nq_x$; the person-years lived by the hypothetical cohort as it ages from age x to age $x + n$, denoted by ${}_nL_x$ (also equivalent to the population aged x, $x + n$ in a "stationary population" experiencing a number of births each year equal to the radix of the life table); and the person-years lived by the hypothetical cohort from age x onward, denoted by $T(x)$.

M

MATCH: Comparison of records for an individual person or housing unit from the census and from the evaluation survey. If the responses agree or fall within a predetermined tolerance, the records are considered a "match."

MATCHING BIAS: Bias in the estimate of the coverage of the census due to incorrect outcomes of the matching process.

MEAN AGE OF CHILDBEARING: Average age at which a mortality-free "cohort" of women bear their children according to a set of "age-specific fertility rates."

MEAN AGE OF CHILDBEARING IN THE POPULATION: Average age of the mothers of the children born in a population during a year. This measure incorporates the effects of both mortality and the age distribution.

MEDIAN: The value associated with the central member of a set that is ordered by size or some other characteristic expressed in numbers.

MIGRATION RATE: Number of migrants during a specified period divided by the person-years lived of the population exposed to migration.

MODEL LIFE TABLE: Expression of typical mortality experience derived from a group of observed "life tables."

MYERS INDEX: An index of digit preference that essentially sums in turn the population ending in each digit over some age range, often 10-89, expressing the total as a percentage of the total population, and which avoids the bias introduced by the fact that the population is not evenly distributed among all ages by repeating the calculations 10 times, once for each beginning digit, and averaging the results. The difference between the average percentage for each digit and the expected value of 10 percent provides a measure of the preference for or avoidance of the digit over the age range considered.

N

NET ERROR: The difference between the census count for any category and the correct count for that category.

NET MIGRATION: The difference between gross immigration and gross emigration.

NON-MATCH: Comparison of records for an individual person or housing unit from the census and from the evaluation survey. If the responses do not agree or do not fall within a predetermined tolerance, the records are considered a "non-match."

NONSAMPLING ERROR: Error from a source other than sampling, such as imperfect selection, bias in response or estimation, errors of observation and recording, data keying and tabulation errors, etc.

O

OUT-OF-SCOPE BIAS: Bias that arises from the erroneous inclusion of out-of-scope persons or housing units among the reports of either of two sources used in the matching operation.

OVERCOUNT: Type of coverage error in which persons or housing units are enumerated (counted) more than once or are enumerated when they should not be.

OWN-CHILDREN METHOD: A refinement of the "reverse-survival" procedure for fertility estimation, whereby estimates of "age-specific fertility rates" for the recent past are obtained by relating mothers to their own children, using information on relationship and other characteristics available from a census or survey.

P

PARAMETER: Unknown quantity to be estimated.

PERIOD FERTILITY: Fertility experienced during a particular period of time by women from all relevant birth or marriage "cohorts."

POPULATION CHANGE DUE TO MIGRATION: The sum of in-migrants minus out-migrants during a specified period of time. The change also may be expressed as a rate by dividing the change by person-years lived in the population during the same period.

R

RADIX: The hypothetical birth "cohort" of a "life table." Common values are 1, 1,000 and 100,000.

RATE OF NATURAL INCREASE: The difference between the births and deaths occurring during a given period divided by the number of person-years lived by the population during the same period. This rate, which specifically excludes changes resulting from migration, is the difference between the "crude birth rate" and the "crude death rate."

RECONCILIATION INTERVIEW: Interview of the same household to resolve discrepancies between the census and the coverage evaluation survey.

RECORD CHECK: Comparison of census data for an individual or household with independent records for the same person or household.

REINTERVIEW SURVEY: Survey in which the household is interviewed a second time, using the same or different questions.

RESPONSE BIAS: Difference between the expected value of the statistic and its true value.

RESPONSE CORRELATION BIAS: Source of bias arising from a failure to maintain independence between two collection systems.

RESPONSE ERROR: Difference between reported answers and correct answers.

RESPONSE VARIANCE: Average of the squared differences between the values of a statistic for the individual repetitions and the expected value of these observations over a large number of repetitions.

REVERSE SURVIVAL: A technique to estimate an earlier population from an observed population, allowing for those members of the population who would have died according to observed or assumed mortality conditions. It is used as a method of estimating fertility by calculating from the observed number of survivors of a given age x the expected number of births that occurred x years earlier. In situations for which both fertility and mortality are known or can be reliably estimated, reverse survival can be used to estimate migration.

ROBUSTNESS: A characteristic of estimates that are not greatly affected by deviations from the assumptions on which the estimation procedure is based.

S

SAMPLING ERROR: The difference between a population value and an estimate thereof, derived from a random sample, which is due to the fact that only a sample of values is observed.

SEX RATIO AT BIRTH: Number of male births for each female birth or male births per 100 female births.

SINGULATE MEAN AGE AT MARRIAGE (SMAM): A measure of the mean age at first marriage, derived from a set of proportions of people single at different ages or in different age groups, usually calculated separately for male and female.

STABLE AGE DISTRIBUTION: Proportion of persons of different age groups in a stable population.

STABLE POPULATION: A population exposed for a long time to constant fertility and mortality rates, and closed to migration, establishes a fixed age distribution and constant growth rate characteristic of the vital rates.

STANDARD ERROR: Measure of dispersion of an estimate from the expected value.

STATIONARY POPULATION: A "stable population" that has a zero growth rate, with constant numbers of births and deaths per year. Its age structure is determined by the mortality rates and is equivalent to the person-years lived $({}_nL_x)$ column of a conventional "life table."

SURVIVAL RATIO: The probability of surviving between one age and another; often computed for age groups, in which case the ratios correspond to those of the person-years lived function, ${}_nL_x$, of a "life-table."

SYNTHETIC PARITY: The average parity calculated for a hypothetical cohort exposed indefinitely to a set of period "age-specific fertility rates."

T

TOTAL ERROR: The sum of the coverage and content errors in the census count for any category.

TOTAL FERTILITY RATE (TFR): The average number of children that would be born per woman if all women lived to the end of their childbearing years and bore children according to a given set of "age-specific fertility rates"; also referred to as total fertility. It is frequently used to compute the consequence of childbearing at the rates currently observed.

TRUE VALUE: Value which producers of a statistic intended to measure.

U

UNITED NATIONS AGE/SEX ACCURACY INDEX: An index of age-reporting accuracy that is based on deviations from the expected regularity of population size and sex ratio, age group by age group. The index is calculated as the sum of: (a) the mean absolute deviation from 100 of the age ratios for males; (b) the mean absolute deviation from 100 of the age ratios for females; and (c) three times the mean of the absolute difference in reported sex ratios from one age group to the next. The United Nations defines age/sex data as "accurate," "inaccurate," or "highly inaccurate," depending upon whether the index is less than 20, from 20 to 40, or greater than 40.

V

VARIANCE: Square of the standard error or average of the squared differences between the values of the statistic on the individual repetitions and the expected value of these observations taken over a large number of repetitions.

W

WHIPPLE'S INDEX: A measure of the quality of age-reporting based on the extent of preference for a particular target digit or digits. The index essentially compares the reported population at ages ending in the target digit or digits with the population expected on the assumption that population is a linear function of age. For a particular age range, often 23-62, the population with ages ending in the target digits is divided by one-tenth of the total population; the result is then multiplied by 100 and divided by the number of different target digits. A value of 100 indicates no preference for those digits, whereas values over 100 indicate positive preference for them.