METHODOLOGICAL REPORTS 1

Demographic
and Health
Surveys

An Assessment of DHS-I Data Quality

Institute for
Resource Development Inc.
A Macro Systems Company

The Demographic and Health Surveys (DHS) is a nine-year project to assist government and private agencies in developing countries to conduct national sample surveys on population and health. DHS is funded by the U.S. Agency for International Development and administered by the Institute for Resource Development.

The main objects of the DHS program are: (1) to provide decisionmakers in the survey countries with data and analyses useful for informed policy choices, (2) to expand the international population and health database, (3) to advance survey methodology, (4) to develop in participating countries the skills and resources necessary to conduct demographic and health surveys.

For information about the Demographic and Health Surveys program, write to DHS, IRD/Macro Systems, 8850 Stanford Boulevard, Suite 4000, Columbia, MD 21045, USA (Telephone 301-290-2800; Telex 87775; Telefax 301-290-2999).

Demographic and Health Surveys
Methodological Reports No. 1

An Assessment of DHS-I Data Quality

Institute for Resource Development/Macro Systems, Inc.
Columbia, Maryland USA

December 1990

The recommended citation for this publication is:

Institute for Resource Development. 1990. *An Assessment of DHS-I Data Quality.* DHS Methodological Reports, No.1. Columbia, Maryland: Institute for Resource Development/ Macro Systems, Inc.

Contents

PREFACE ix

INTRODUCTION 1

ASSESSMENT OF THE QUALITY OF DATA USED TO ASCERTAIN ELIGIBILITY AND AGE IN THE DEMOGRAPHIC AND HEALTH SURVEYS
by Shea Oscar Rutstein and George T. Bicego

1 **Introduction** 5

2 **Quality of Data Used to Determine Eligibility for the Individual Interview** 7
2.1 Household Interview Nonresponse 7
2.2 Age Reporting in the Household Questionnaire 7
2.3 Reporting of Household Residency 13
2.4 Simulating the Effects of Exclusion 21

3 **Age Reporting in the Individual Interview** 25
3.1 Reported Age vs. Imputed Age 25
3.2 Digit Preference 25
3.3 Five-year Age Group Distortions 27
3.4 Comparison with the WFS Surveys 34

4 **Summary and Conclusions** 35
4.1 Eligibility 35
4.2 Response Rates 35
4.3 Age of Respondent 35
4.4 Recommendations 36

References 37

ASSESSMENT OF THE QUALITY OF DATA ON AGE AT FIRST SEXUAL INTERCOURSE, AGE AT FIRST MARRIAGE, AND AGE AT FIRST BIRTH IN THE DEMOGRAPHIC AND HEALTH SURVEYS
by Ann K. Blanc and Naomi Rutenberg

1 **Introduction** 41

2 **Date of First Union** 43
2.1 Completeness of Information 43
2.2 Heaping 47
2.3 Trends in the Median Age at Marriage 48
2.4 Proportion of Women Ever Married at Ages 15-19 and 20-24 51

2.5 Comparison of DHS Surveys and Other Sources 53

3 **Date of First Birth** 57
3.1 Completeness of Information 57
3.2 Heaping 60
3.3 Trends in the Median Age at First Birth 61
3.4 Proportion of Women Who Had a First Birth at Ages 15-19 and 20-24 64
3.5 Interval Between Date of First Union and Date of First Birth 65

4 **Age at First Sexual Intercourse** 67
4.1 Completeness of Information 67
4.2 Trends in the Median Age at First Sexual Intercourse 67
4.3 Interval Between Age at First Sexual Intercourse and First Union 67
4.4 Interval Between Age at First Sexual Intercourse and First Birth 70

5 **Summary and Conclusions** 71

References 79

ASSESSMENT OF THE QUALITY OF BIRTH HISTORY DATA IN THE DEMOGRAPHIC AND HEALTH SURVEYS

by Fred Arnold

1 **Introduction** 83

2 **Completeness of Reporting of Children's Birth Dates** 84

3 **Displacement of Children's Birth Dates** 89
3.1 Description of the Problem 89
3.2 Evidence for the Displacement of Birth Dates 90
3.3 Implications 92
3.4 Possible Solutions to the Problem 93

4 **Age Heaping** 95

5 **Miscalculation of Year of Birth** 97
5.1 Description of the Problem 97
5.2 Implications 98
5.3 Possible Solutions to the Problem 99

6 **Coverage of Live Births** 100
6.1 Age at First Birth 100
6.2 Sex Ratios at Birth 101

7 **Comparison of DHS and WFS Fertility Rates** . . 104

8 **Summary and Conclusions** 108

References 110

ASSESSMENT OF THE QUALITY OF DATA USED FOR THE DIRECT ESTIMATION OF INFANT AND CHILD MORTALITY IN THE DEMOGRAPHIC AND HEALTH SURVEYS

by Jeremiah M. Sullivan, George T. Bicego, and Shea Oscar Rutstein

1 **Introduction** 115

2 **Collection of DHS Mortality Data** 116
2.1 Questionnaires 116
2.2 Field Procedures 116
2.3 Data Processing 116

3 **Date of Birth Data** 117
3.1 Completeness of the Data 117
3.2 Accuracy of the Data 118
3.3 Impact on Estimates of Infant Mortality . . . 118
3.4 Data Entry and Imputation Programs 119

4 **Age at Death Data** 121
4.1 Completeness of the Data 121
4.2 Accuracy of the Data 122
4.3 Impact on Mortality Estimates 123

5 **Completeness of Event Reporting** 126
5.1 Internal Consistency 126
5.2 Comparison of DHS Data with Other Sources 132

6 **Summary and Recommendations** 135
6.1 Quality of the Reported Data 135
6.2 Recommendations 135

References 137

Appendix
Birth History Section of the DHS-I Questionnaire (Section 2. Reproduction) 139

SUMMARY OF DHS-I SURVEYS, 1985-1990 141

Tables

ASSESSMENT OF THE QUALITY OF DATA USED TO ASCERTAIN ELIGIBILITY AND AGE IN THE DEMOGRAPHIC AND HEALTH SURVEYS

2.1 Household nonresponse rates, household age and sex ratios, and Myers and United Nations indices for household data 8

2.2 Indices of age eligibility distortion based on household data from DHS and WFS surveys and national censuses 13

2.3 (A) Number of women reported to reside in interviewed households, (B) the percentage of resident women not sleeping in the household during the night before the survey, and (C) the standardized percentage of nonresident women sleeping in the household during the night before the survey, by five-year age group 14

2.4 Estimated percentage of women who were excluded from eligibility for the individual interview due to mis-recording of "sleeping away" status, by age group 16

2.5 Difference in the estimated percentage of women who were excluded from eligibility due to misrecording of "sleeping away" status, between age groups immediately inside and outside the age eligibility boundaries 20

2.6 Results of simulations to estimate the effect of lower boundary, upper boundary, and "sleeping away" exclusions on total fertility rate, under-five mortality, and contraceptive prevalence 23

2.7 Percent distribution of eligible women by response to the individual interview 24

3.1 Percent distribution of respondents to the individual questionnaire according to information given and imputed for date of birth and current age 26

3.2 Percent distribution of women 20-49 by reported terminal digit of age (individual questionnaire) and Myers index 27

3.3 Percent distribution of women by five-year age group (individual questionnaire) 28

3.4 Age Ratios for five-year age groups, Demographic and Health Surveys, 1986-1989, and comparison of the age-group distortion index for DHS and WFS surveys 34

ASSESSMENT OF THE QUALITY OF DATA ON AGE AT FIRST SEXUAL INTERCOURSE, AGE AT FIRST MARRIAGE, AND AGE AT FIRST BIRTH IN THE DEMOGRAPHIC AND HEALTH SURVEYS

1.1 Background information and type of data collected on union status, date of first union, and age at first union 42

2.1 Percent distribution of ever-married women by completeness of information on date of first union, by age group 44

2.2 Indices of heaping on year of first union and duration since first union ending in 0 or 5 . . . 47

2.3 Median age at first union, by age of woman at the time of the survey 48

2.4 Percentage of women ever married at ages 15-19 and 20-24, for selected years prior to the survey 52

2.5 Comparison of the percentage of women ever married in DHS surveys with the percentage ever married in previous censuses or surveys, by age at the time of the previous census or survey 54

3.1 Percent distribution of women who have ever given birth by completeness of information on date of first birth, by age group 57

3.2 Indices of heaping on year of first birth and duration since first birth ending in 0 or 5 . . . 60

3.3 Median age at first birth, by age at the time of the survey 61

3.4 Percentage of women who ever had a birth at ages 15-19 and 20-24, for selected years prior to the survey 65

3.5 Percent distribution of the interval between first union and first birth 66

4.1 Median age at first sexual intercourse, by age of woman at the time of the survey 68

4.2 Percentage distribution of the interval between first sexual intercourse and first union 69

4.3 Percent distribution of women who have ever given birth by the interval between first sexual intercourse and first birth 70

5.1 Summary of data quality indicators 72

ASSESSMENT OF THE QUALITY OF BIRTH HISTORY DATA IN THE DEMOGRAPHIC AND HEALTH SURVEYS

2.1 Percent distribution of children born to survey respondents by completeness of information on date of birth and age of children 85

2.2 Percentage of children with complete information on year and month of birth by number of years since birth of child 86

2.3 Percentage of children with complete information on year and month of birth by age and education of mother 87

2.4 Percentage of children with complete information on year and month of birth by selected background characteristics 88

3.1 Number of questions in the DHS-I and DHS-II core questionnaires that depend on children's year of birth 89

3.2 Number of births by calendar year 90

3.3 Birth year ratios by survival status of births . . 91

4.1 Age ratios for living children by single year of age 96

5.1 Percentage of children whose month of birth falls in the month of interview or earlier 98

6.1 Average number of children ever born by age of woman 101

6.2 Median age at first birth by age of woman at the time of the survey 102

6.3 Sex ratios at birth for five-year periods prior to the survey 103

ASSESSMENT OF THE QUALITY OF DATA USED FOR THE DIRECT ESTIMATION OF INFANT AND CHILD MORTALITY IN THE DEMOGRAPHIC AND HEALTH SURVEYS

3.1 Percentage of births during the fifteen-year calendar period preceding the survey with incomplete information on date of birth by survival status 117

3.2 Birth ratios for the fifth calendar year preceding the survey by survival status 119

3.3 Births in the five years preceding the survey, number of living and dead children, and infant mortality rates, based on standard recode data and reimputed data 120

4.1 Percentage of deaths among children under five with incomplete information on age at death by calendar period in which the death occurred 121

4.2 Total reported deaths among children under five and the number of deaths with incomplete age at death information by type of defect in information 122

4.3 Index of heaping of deaths at twelve months of age by calendar period in which the death occurred 123

4.4 Estimates of infant and child mortality for the ten-year period preceding the survey, adjusted for heaping of deaths at twelve months of age 125

5.1 Infant mortality rates and ratio of neonatal to infant mortality by calendar year period preceding the survey 126

5.2 Neonatal mortality rates and the ratio of deaths under seven days to all neonatal deaths by calendar year period preceding the survey 127

5.3 Infant mortality rates by sex of child and sex risk ratio (M/F) for the ten-year period preceding the survey 128

5.4 Infant mortality rates by birth order of child, for the ten-year calendar period preceding the survey 129

5.5 Infant mortality rates by age of mother at birth of child, for the ten-year calendar period preceding the survey 130

5.6 Infant mortality rates by length of retrospective birth interval for the ten-year calendar period preceding the survey 131

5.7 Infant mortality rates by mother's education for the ten-year calendar period preceding the survey 132

5.8 Estimates of under-five mortality for a fixed five-year reference period, WFS and DHS Surveys 133

5.9 Estimates of infant and child mortality for a fixed five-year reference period, WFS and DHS surveys 134

Figures

ASSESSMENT OF THE QUALITY OF DATA USED TO ASCERTAIN ELIGIBILITY AND AGE IN THE DEMOGRAPHIC AND HEALTH SURVEYS

2.1 Single-year age distribution of women 10 or more years of age (household data), Demographic and Health Surveys in Ghana (1988) and Brazil (1986) 9

2.2 Age distribution of household members by sex, Demographic and Health Surveys in Ghana (1988) and Brazil (1986) 10

2.3 Percentage of women not reported as sleeping in a household during the night before the survey by age, Demographic and Health Surveys, 1986-1989 17

3.1 Age distribution of women based on data from the individual questionnaire, Demographic and Health Surveys, 1986-1989 . . 29

ASSESSMENT OF THE QUALITY OF DATA ON AGE AT FIRST SEXUAL INTERCOURSE, AGE AT FIRST MARRIAGE, AND AGE AT FIRST BIRTH IN THE DEMOGRAPHIC AND HEALTH SURVEYS

2.1 Median age at first union, selected WFS and DHS surveys, by age at the time of the DHS survey 49

3.1 Median age at first birth, selected WFS and DHS surveys, by age at the time of the DHS survey 62

5.1 Median age at first sexual intercourse, first union, and first birth, Demographic and Health Surveys, 1986-1989 73

ASSESSMENT OF THE QUALITY OF BIRTH HISTORY DATA IN THE DEMOGRAPHIC AND HEALTH SURVEYS

3.1 Number of births by calendar year, selected Demographic and Health Surveys, 1986-1987 . . 92

3.2 Average succeeding birth interval by year of birth, selected Demographic and Health Surveys, 1986-1987 94

7.1 Trends in the total fertility rate (TFR) for women 15-34, selected DHS and WFS surveys 105

ASSESSMENT OF THE QUALITY OF DATA USED FOR THE DIRECT ESTIMATION OF INFANT AND CHILD MORTALITY IN THE DEMOGRAPHIC AND HEALTH SURVEYS

4.1 Age at death for children who died at less than two years of age, selected Demographic and Health Surveys, 1987 124

Preface

One of the objectives of the Demographic and Health Surveys (DHS) program is to advance the methodology and procedures pertaining to national-level surveys in the fields of population and maternal and child health.

In the course of implementing 34 surveys during the first five-year phase of the project and the initiation of 25 surveys during the second phase, questions and issues have arisen regarding the design and implementation of these surveys.

The purpose of the *DHS Methodological Reports* series is to examine some of these questions and issues and to provide answers, explanations, and solutions which will be of benefit to survey researchers, particularly those in developing countries.

Survey methodology can have a substantial impact on data quality. This report deals with issues of data quality in the DHS surveys. The major objective of this report is to provide data users with a global assessment of the quality of DHS data. A secondary objective is to examine particular problems of data quality and to suggest reasons for their occurrence.

Future reports in this series will examine sampling, survey implementation, and analysis issues, with a view to improving survey research efforts in the future.

Martin Vaessen
Project Director

Introduction

The mandate of the Demographic and Health Surveys (DHS) program is to conduct national sample surveys in developing countries. The program, which covers the period 1984-1993, is divided into two 5-year phases, with 1989 being an overlap year. The overlap year was designed into the program to allow a period for reviewing DHS survey procedures without causing a break in the initiation of new surveys.

The review of DHS procedures consisted of three components: 1) an assessment of survey instruments and reports which was based on the experience of host country institutions and DHS project staff, 2) an assessment of the need for population and health data by developing country institutions, the U.S. Agency for International Development (USAID), and international agencies, and 3) an assessment of the quality of the DHS data. The results of the third component are reported in this volume.

Objectives

The objectives of this assessment are to evaluate the quality of the DHS data and to determine if there is a need to modify questionnaires and field procedures in future surveys. The assessment focuses on data problems which derive from nonsampling error (e.g., respondent underreporting of events, interviewer misrecording of information, errors arising from questionnaire design, etc.). The issue of sampling error will be examined in a separate report.

DHS surveys include a household questionnaire, in which members and visitors in sampled households are listed, and an individual questionnaire in which information reported by women of childbearing age (usually 15-49 years) is recorded. The individual questionnaire includes sections on the respondent's characteristics, her birth history, knowledge and use of contraception, and maternal and child health.

The data from 22 surveys are evaluated in this assessment (see table). These include all national-level surveys for which a standard recode data file was available as of mid-1989. The regional breakdown of surveys is: sub-Saharan Africa (9), North Africa (2), Asia (3), Latin America and the Caribbean (8). For surveys which are not self-weighting, the analysis presented in this report is based on weighted data.

The evaluation is comparative in nature. The data from each survey are subjected to a series of tests, which differ depending on the type of data under investigation. To the extent feasible, the following aspects of data quality are considered:

- the completeness of the recorded data,
- the accuracy of the data,
- the consistency of the data relative to expected patterns, and
- a comparison of estimates based on DHS data with estimates from other sources.

In addition to documenting the extent of specific data problems, simulation analysis is used to show the potential impact of the data errors on estimates of fertility, mortality, and contraceptive prevalence. The report also includes recommendations for reducing errors in future surveys.

Surveys included in the assessment of DHS data quality

Country	Year of Survey	Sample Size	Respondents	Sample Structure
SUB-SAHARAN AFRICA				
Botswana	1988	4368	AW 15-49	Weighted
Burundi	1987	3970	AW 15-49	Weighted
Ghana	1988	4488	AW 15-49	Self-Weighting
Liberia	1986	5239	AW 15-49	Weighted
Mali	1987	3200	AW 15-49	Weighted
Senegal	1986	4415	AW 15-49	Self-Weighting
Togo	1988	3360	AW 15-49	Self-Weighting
Uganda	1988/89	4730	AW 15-49	Weighted
Zimbabwe	1988/89	4201	AW 15-49	Self-Weighting
NORTH AFRICA				
Morocco	1987	5982	EMW 15-49	Self-Weighting
Tunisia	1988	4184	EMW 15-49	Self-Weighting
ASIA				
Indonesia	1987	11884	EMW 15-49	Weighted
Sri Lanka	1987	5865	EMW 15-49	Weighted
Thailand	1987	6775	EMW 15-49	Weighted
LATIN AMERICA/ CARIBBEAN				
Brazil	1986	5892	AW 15-44	Weighted
Colombia	1986	5329	AW 15-49	Weighted
Dominican Rep.	1986	7649	AW 15-49	Weighted
Ecuador	1987	4713	AW 15-49	Self-Weighting
Guatemala	1987	5160	AW 15-44	Self-Weighting
Mexico	1987	9310	AW 15-49	Weighted
Peru	1986	4999	AW 15-49	Self-Weighting
Trinidad & Tob.	1987	3806	AW 15-49	Self-Weighting

AW = All Women
EMW = Ever-Married

Findings

The results of the data quality assessment are presented in four papers, each reporting on a particular type of data. The four papers must be read in their entirety to understand the complexities of the data evaluation task and to understand the quality of the DHS data. Nevertheless, it seems appropriate to summarize the principal findings in this introduction.

The first paper evaluates the age and residence data from the household questionnaires (which determine eligibility for the individual questionnaire) and the age data for women interviewed with the individual questionnaire. Analysis of the household data suggests some intentional misrecording of information in order to make some women ineligible for the individual questionnaire. The evidence of exclusion is greater for surveys in sub-Saharan Africa than for surveys in other regions. Simulation analysis indicates that, under extreme assumptions about the characteristics of excluded women, the total fertility rate could be overestimated by about 4 percent in some sub-Saharan surveys and by about 2 percent in other regions. On the other hand, under-five mortality could be underestimated by about 4 percent in sub-Saharan Africa and by about 2 percent elsewhere. Estimates of contraceptive prevalence could be little affected even under extreme assumptions of exclusion.

The second paper reviews data on three aspects of the early stages of the reproductive period—age at first sexual intercourse, age at first union, and age at first birth. A significant finding from the analysis is the utility of age at first sexual intercourse as an indicator of the beginning of exposure to the risk of pregnancy. Age at first sexual intercourse was reported relatively completely, despite concerns about collecting such sensitive data, and often appeared to be more reliably reported than date of first union.

The third and fourth papers examine the birth history data from the individual questionnaire and the impact of data error on the estimation of fertility rates and infant and child mortality rates. Various problems with the birth history data were found, but the probable effects on fertility and mortality estimates were judged to be minimal in most surveys. An important part of the evaluation was the comparison of estimates based on DHS data with those from other data sources. DHS fertility and mortality estimates for time periods about a decade before the survey were in close agreement with estimates from the World Fertility Survey. In the case of fertility estimates, the agreement was especially close.

The evaluation of the birth history data paid particular attention to the problem of systematic displacement of births in time. The individual questionnaire used in the DHS surveys is considerably longer than the questionnaires used in the World Fertility Survey or the Contraceptive Prevalence Surveys, due to the addition of several pages of health questions applicable to children under five. In some surveys it appears that interviewers misrecorded the birth dates of some children in order to avoid asking the health questions. Little evidence of birth displacement was found in the surveys from Asia and Latin America and the Caribbean, but it was evident in most of the African surveys. Birth displacement was found to have little effect on multiple-year estimates of the level of fertility. However, the effect on estimates of fertility trends can be greater because systematic birth displacement can affect estimated rates for different time periods in opposite directions.

Overall, the assessment of the quality of the DHS data revealed problems of the type which are typically found in retrospective surveys. However, gross errors that would seriously affect demographic estimates were not detected. Estimates of contraceptive prevalence (at the time of a survey) and estimates of fertility and childhood mortality (at the time of a survey and extending back to as far as fifteen years preceding a survey) appeared to be reliable in most surveys. The data problems that were identified were most severe in the surveys in sub-Saharan Africa. Based on this assessment, some changes are being made in the DHS questionnaires and field procedures. Although data quality will always be an issue in retrospective surveys, it is expected that the changes resulting from this assessment will reduce data errors in the future.

Assessment of the Quality of Data Used to Ascertain Eligibility and Age in the Demographic and Health Surveys

SHEA OSCAR RUTSTEIN
GEORGE T. BICEGO

Acknowledgments

The authors are grateful for the many helpful comments and suggestions provided by Kenneth Hill, Anne Cross, Martin Vaessen, Noreen Goldman, and Jeremiah Sullivan. David Cantor provided valuable programming assitance.

1 Introduction

The purpose of this report is to evaluate the data from the Demographic and Health Surveys (DHS) in two areas: eligibility of women for the individual questionnaire and quality of information on age of respondent. Both are crucial to an assessment of the general quality of DHS data, since omission of certain women from the individual interview could potentially bias almost all the results obtained and deficiencies in age data could similarly affect most of the results. Age is a particularly important variable, since many of the estimates derived from these surveys, especially fertility rates, are heavily age-dependent. This report is modelled on one that was produced for the World Fertility Survey (see Rutstein, 1985), the data from which are compared to those used here.

DHS surveys attempt to interview a representative sample of women in the childbearing ages (generally 15-49). Most surveys include all women in these ages, while some (mostly in Asia and North Africa) include only those who have ever married. Some women are excluded for logistical reasons, such as those who live in institutions (hospitals, military barracks, boarding schools) and those with no fixed residence (vagrants, nomads). In order to avoid over- or under-counting, each woman who is in the proper age range must be associated with one and only one household. One way to accomplish this is to interview all women of childbearing age who usually live in the selected household (de jure rule). Another is to interview all women of childbearing age who spent the night before the interview in the selected household (de facto rule). Since the latter rule should logically increase response rates (women are more likely to be at home if they were there the previous night), all but two DHS surveys have adopted the de facto eligibility rule.

In addition to not living in a private household, there are various other ways that a woman who should be eligible for the individual interview can be excluded. She can live in a household that was not interviewed because no one was at home on any of the occasions when the interviewer came or because someone in the household refused to grant the interview. Or, her age as recorded on the household questionnaire could be misreported, such that she erroneously falls out of the age range for eligibility. Alternatively, she could be excluded from the individual interview because she was misreported as not having spent the previous night in the household. Finally, even if she is correctly classified as eligible, she may not be interviewed individually either because she is not at home, refuses, is ill or deaf, etc. Of course it is also possible for ineligible women to be included in the individual interview for any of the above reasons (households not selected for the sample being mistakenly interviewed, women outside the age range being reported as within the age range, etc.), however, given the effort involved in conducting the individual interview, it would seem that there is much more incentive to falsely omit eligible women than to falsely include women who are ineligible.

Thus, in order to evaluate the quality of the data used for determining eligibility for the individual questionnaire, the following are examined:

- Household response rates,
- Age reporting in the household questionnaire and exclusion from the individual questionnaire due to misreporting of age,
- Reporting of residence and previous night's location, and
- Individual response rates.

Results from the above analyses provide a rough idea of the possible extent of omission of eligible women, however, it tells little about the potential bias in the results obtained from those who were interviewed. Bias arises when women who were excluded from the survey differ from those who were interviewed in terms of the variables measured. A simulation exercise was carried out to examine the potential effects of exclusion of eligible women on the estimation of various demographic parameters (total fertility rate, under-five mortality, and contraceptive prevalence rate). The results presented here indicate the potential level of bias associated with errors in data.

While ages reported during the household interview are used in determining eligibility for the individual interview, it is the subsequent face-to-face interviews with eligible women that provide the data on age and birth dates that are the basis of most demographic analyses. The value of these analyses depends to a large extent on the accuracy of age reporting by respondents. For example, classifying even a relatively small number of respondents into the wrong five-year age group can affect the estimates of age-specific fertility rates, average number of children ever born by age, and proportions married by age—all basic demographic measures. Depending on whether dates of other events such as marriage and births are estimated independently from age of respondent, age misreporting can affect data on age at marriage and age at birth.[1] To evaluate the quality of these age data, the following are examined:

[1]For a more comprehensive discussion of the types of biases that can result from misreporting of age, see Rutstein 1985:12-13.

- Patterns of age reporting and age imputation,
- Digit preference, and
- Gross misstatement of age.

This report is divided into three sections. The first covers household response rates, the quality of data used to determine eligibility for individual interview (namely, age and residency the night before the interview as reported in the household questionnaire), and individual response rates. The second section is concerned with the quality of age reporting in the individual questionnaire. The third section consists of summary and conclusions.

2 Quality of Data Used to Determine Eligibility for the Individual Interview

2.1 HOUSEHOLD INTERVIEW NONRESPONSE

DHS samples are designed to produce a representative sample of households. The sampling process used in the DHS surveys is described in the DHS *Sampling Manual* (Institute for Resource Development, 1987). Generally, the samples are selected in two stages. The first stage involves selecting geographic units of a known size (often census enumeration areas). All of the households in these selected areas are then listed and the required number of households is selected from each unit. While problems of representativeness of the sample can appear in the frame of geographical units, in identifying the selected units, and in listing all the households within selected units, this paper covers only the problems of nonresponse among certain households.

After the selected households have been identified, a list is made of all persons who are usual members of the household (including domestic employees), as well as all current visitors. From a responsible adult in each household, information is obtained on the age and sex of each person and whether or not he/she spent the previous night in the dwelling. This information is then used to determine eligibility for the individual interview.

In most DHS surveys, women are eligible for the individual interview if they are age 15 to 49 and slept in the dwelling during the previous night (de facto residence). In two surveys, Brazil and Guatemala, the upper boundary of age eligibility was 44 years. In Brazil and Indonesia, a de jure residence standard was used, so that only usual household members were considered eligible. In five surveys (Indonesia, Morocco, Sri Lanka, Thailand, and Tunisia), only women who had ever been married were eligible for the individual questionnaire.

Response rates can have important effects on data quality. Households which were part of the original sample, but from which no information was obtained, may differ from households where questionnaires were completed. If these differences involve the parameters being measured, and if the level of household nonresponse is high, the possibility of sampling bias must be considered.

Eighteen of the 22 DHS surveys analyzed here (82 percent) had household nonresponse rates of less than 5 percent (see Table 2.1).[2] In the two countries which substantially exceed this limit, Liberia (11.6 percent) and Togo (7.5 percent), the majority of cases of nonresponse were due to refusal to complete the interview. Unfortunately, there is no information on how the households which were interviewed differ from the households that were not.

2.2 AGE REPORTING IN THE HOUSEHOLD QUESTIONNAIRE

Aside from gender, age is the most important criterion used to determine eligibility for the individual questionnaire. Information on age was collected during the household interview by asking the informant about age in completed years, using the question "How old is he/she?" Since any responsible adult could act as the household informant, proxy reports of age occurred for many eligible women. For a variety of reasons, such proxy reports are likely to be inaccurate, particularly in regard to nonfamily household members. The extent of proxy reporting cannot be determined, however, since the household informant was not identified on the questionnaire.

There are several ways to assess the quality of age data. One common approach is to identify age structures that differ from expected patterns and that are more plausibly explained by misreporting than by real phenomena. For the DHS surveys, standard indices that capture distortions in age data have been computed and compared cross-nationally; where possible, they have also been compared with indices computed for WFS surveys and various national censuses.

[2] Only 77 percent of WFS surveys (30 of 39 surveys) met the same criterion.

Table 2.1 Household nonresponse rates, household age and sex ratios, and Myers and United Nations indices for household data, Demographic and Health Surveys, 1986-1989

Country	Percentage of Households Not Responding[a]	Myers Index[b]	UN Index	Age Ratios[b]				Sex Ratios			
				10-14	15-19	45-49	50-54	10-14	15-19	45-49	50-54
SUB-SAHARAN AFRICA											
Botswana	3.2	8.6	70.8	126	76	73	152	85	102	107	46
Burundi	0.4	12.4	67.6	104	83	85	147	92	122	95	63
Ghana	2.2	30.6	64.0	103	82	98	134	104	111	97	66
Liberia	11.6	24.4	89.6	100	93	78	181	99	97	122	60
Mali	0.2	19.8	70.4	103	79	84	130	107	109	121	80
Senegal	0.5	10.8	58.3	95	94	80	147	96	96	103	61
Togo	7.5	19.6	65.9	113	80	72	156	98	124	85	46
Uganda	1.2	25.2	88.8	103	95	79	148	93	84	149	75
Zimbabwe	5.3	10.8	40.3	101	96	92	116	100	112	105	87
NORTH AFRICA											
Morocco	2.8	13.8	35.4	101	105	101	114	104	97	84	77
Tunisia	2.3	11.6	48.9	100	98	72	143	103	98	122	89
ASIA											
Indonesia	3.5	17.2	34.6	105	100	103	113	100	99	95	85
Sri Lanka	2.2	21.4	44.4	99	102	76	130	104	97	118	81
Thailand	1.5	12.2	31.3	106	106	92	116	106	102	99	89
LATIN AMERICA/CARIBBEAN											
Brazil[c]	2.9	4.4	38.8	94	106	108	85	102	83	89	119
Colombia	4.1	15.0	36.3	89	106	98	105	107	93	98	89
Dominican Republic	2.8	12.2	36.7	99	110	94	120	104	93	96	95
Ecuador	1.5	13.4	65.5	107	94	66	182	100	102	126	71
Guatemala[c]	3.9	13.8	50.6	103	93	74	128	101	109	108	77
Mexico	2.2	14.8	33.5	110	103	89	109	98	100	102	90
Peru	4.3	11.6	37.0	115	89	87	121	99	111	106	93
Trinidad & Tobago	5.7	6.2	56.1	98	92	76	140	102	106	116	79

[a]"Household not present night before interview"; dwellings that were vacant or unoccupied, and dwellings that were destroyed are excluded from calculation of the nonresponse rate.
[b]Females only; for Myers, refers to age 10-69
[c]Based on age groups 40-44 and 45-49

Digit Preference and the Myers Index

In most societies, there is a tendency to report ages ending in certain preferred digits (often zero or five). This phenomenon is called *age heaping*. An index developed by Myers (see Shryock and Siegel, 1971: 206-208) measures the amount of preference for ages ending in certain digits by comparing the proportions ending in each of the ten digits with the expected proportion of 10 percent. In the calculations used here the theoretical limits are zero (no heaping at all) and 180 (all heaped on a single digit). The Myers index for the female populations (10-69 years old) recorded on the household questionnaires in DHS surveys are presented in Table 2.1.

Surveys with values of less than 10 have been classified as those having a low level of digit preference, those with values 10 to 20 as having a moderate level of digit preference, and those with values above 20 as having a high level of digit preference. In the table below, the surveys are grouped according to this rating scheme and within groups are ranked according to increasing score on the Myers index (M).

Most of the surveys fall into the category of moderate age heaping. There is some tendency for the countries in sub-Saharan Africa to fall on the high side and those in Latin America to fall on the low side. The three countries with the highest level of age heaping are all in sub-Saharan Africa.

The single-year age distributions for Ghana and Brazil are presented in Figure 2.1 as examples of high and low tendencies for age heaping.

Prevalence of age heaping based on the Myers index (M) among females listed on the household questionnaires, 22 DHS Surveys

Low (M <10)	Moderate (M 10-20)	High (M >20)
Brazil	Senegal	Sri Lanka
Trinidad & Tobago	Zimbabwe	Liberia
Botswana	Peru	Uganda
	Tunisia	Ghana
	Dominican Republic	
	Thailand	
	Burundi	
	Ecuador	
	Guatemala	
	Morocco	
	Mexico	
	Colombia	
	Indonesia	
	Togo	
	Mali	

Age Group Distortion and the United Nations Index

The United Nations (UN) index[3] is commonly used to assess the amount of distortion in five-year age distributions due to gross misstatement of age and nonresponse. The index measures—by means of a summary score—departures from the expected trend in the size of age groups (i.e., a smooth decline in the numbers of women with increasing age) and changes in sex ratios by age group. The higher the score, the greater the likelihood that age reporting is inaccurate. (Unlike the Myers index, as applied here, computation of the UN index includes both the male and female populations.)

Seven of nine surveys in sub-Saharan Africa and one of eight in Latin America have scores exceeding 60 on the UN index, suggesting poor age reporting. In order of decreasing score they are: Liberia, Uganda, Botswana, Mali, Burundi, Togo, Ecuador, and Ghana (see Table 2.1).

In addition, Senegal and Trinidad and Tobago have scores close to 60. Among the sub-Saharan countries, Zimbabwe has the most plausible age-sex structure. The household age distributions by sex for Ghana and Brazil are presented in Figure 2.2 as examples of surveys with relatively high and relatively low scores on the UN index.

[3] A description of the UN index–formally called the age-sex composition index–is given in **Manual II: Methods of Appraisal of Quality of Basic Data for Population Estimates** (United Nations, 1967).

Figure 2.1 Single-year age distribution of women 10 or more years of age (household data), Demographic and Health Surveys in Ghana (1988) and Brazil (1986)

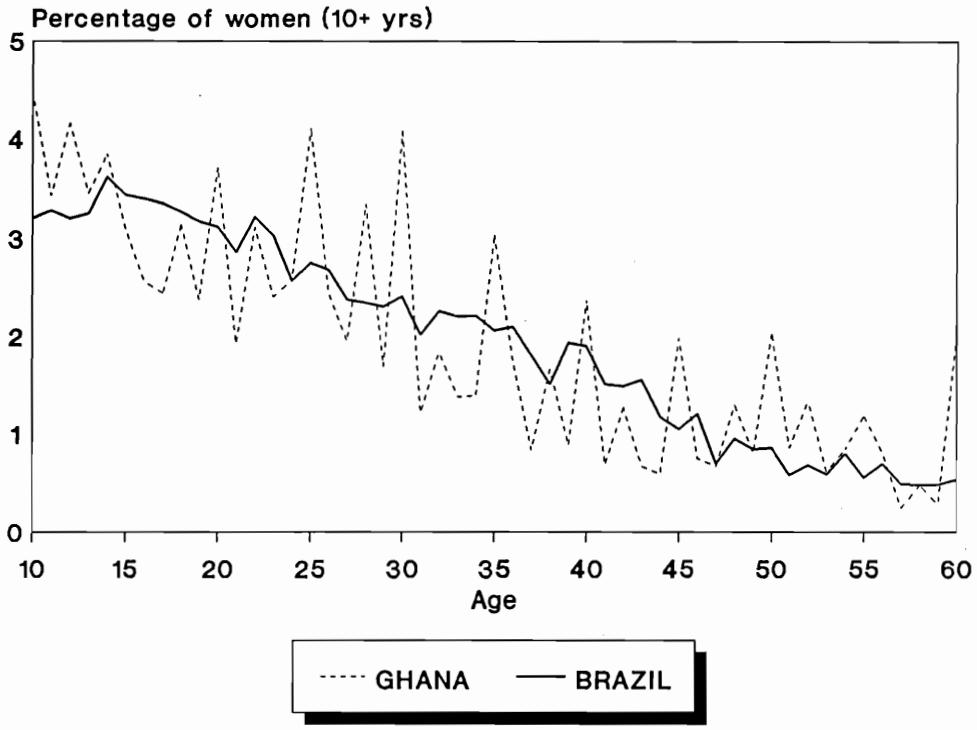
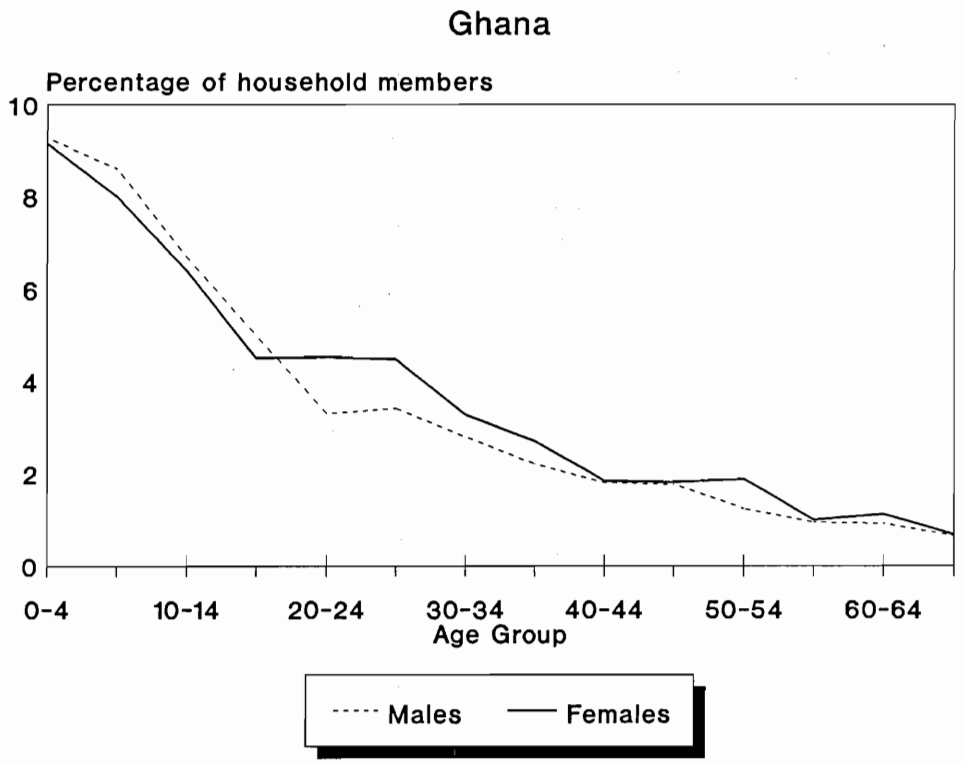
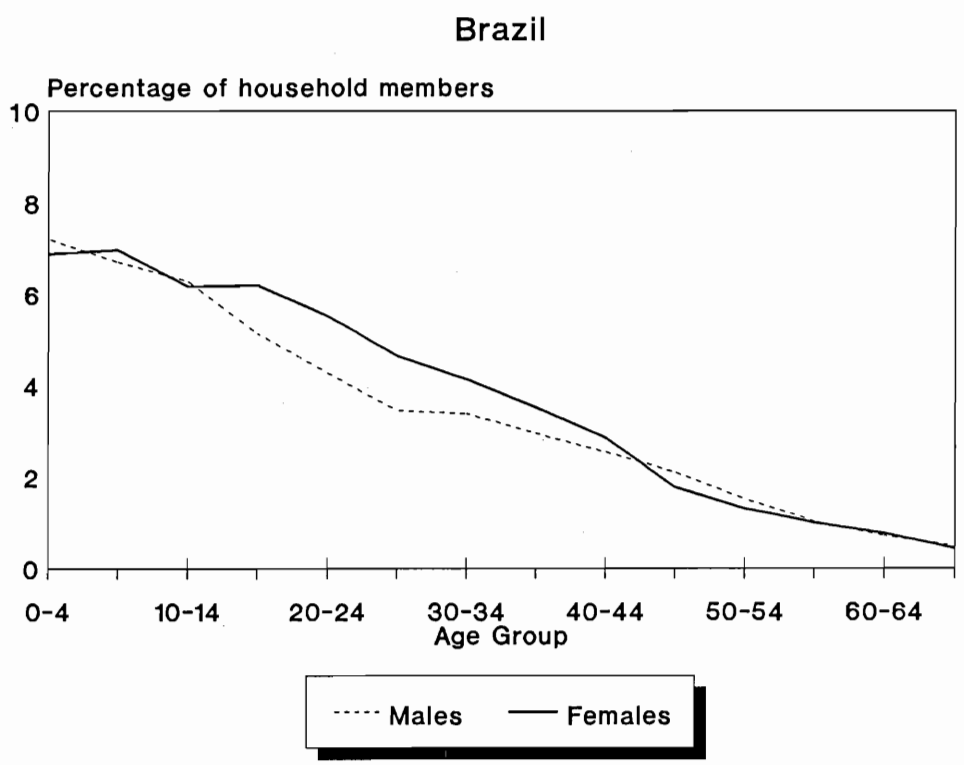


Figure 2.2 Age distribution of household members by sex, Demographic and Health Surveys in Ghana (1988) and Brazil (1986)





Eligible-Age Boundary Effects

The misreporting of women's ages can lead to their exclusion from the individual interview by pushing them out of the eligible age range. DHS data indicate that age misreporting occurs more frequently for age groups at the age eligibility boundaries than for other age groups. One possible explanation is that when interviewers are responsible for both the household and individual questionnaires, those interviewers who want to reduce their workload may push women—especially those whose exact ages are not known—out of the eligible age range in order to reduce the number of women that have to be interviewed. Also, the household informant may exclude some eligible women from being interviewed for personal reasons (e.g., teenage daughters).

There is probably greater incentive as well as greater opportunity for misreporting the ages of women who are near the upper limit of age eligibility. Older women, who generally have less education, often do not know their exact ages. In these cases, the interviewer must spend extra time probing for age information. Also, since older women have had more children than younger women, and more of their children will have died, their interviews must necessarily include lengthy and complicated birth histories. In addition, older women are less likely to recall the dates of important events such as marriage. Women at the lower limit of age eligibility can presumably be interviewed more easily and in a shorter amount of time. However, the more common absence of younger women—e.g., students, unmarried women, and employed women—at the time of the interviewer's first visit, may be an incentive for interviewers to misclassify their ages in order to avoid returning another time for the interview. In addition, some interviewers may be embarrassed to interview young women (15-16 years), who are clearly still adolescents and who for the most part are not yet sexually active.

Examination of age group ratios (for women) and sex ratios for the age groups immediately above and below the age eligibility boundaries can help identify whether such systematic exclusion of eligible women (or, less likely, inclusion of ineligible women) has occurred.[4] In Table 2.1, age and sex ratios are presented for the age groups bordering the upper and lower limits of age eligibility. If women have been systematically excluded from the eligible population, the age ratios for the eligible age group (immediately inside the boundary) would be low relative to those for the ineligible age group (immediately outside the boundary). The opposite would be true of the sex ratios for these age groups.

Upper Boundary Effect. Age ratios for the DHS surveys give some indication of a boundary effect at the upper limit of age eligibility. Eighteen of 22 surveys show strong evidence of exclusion and three surveys (Indonesia, Morocco, and Colombia) show a more moderate effect. In only one survey, Brazil, is there evidence of a downward transfer of women into the oldest eligible age group.[5] The apparent out-transference of eligible women is most pronounced in sub-Saharan countries (particularly, Botswana, Liberia, and Togo), but is very strong in Ecuador as well.

Differences in the sex ratios for the 45-49 and 50-54 age groups indicate that almost all surveys may have suffered from the out-transference of eligible women. Sixteen surveys show large differences in the sex ratios, and four (Morocco, Indonesia, Thailand and Colombia) show modest differences of between 5 and 10 points. Only in the Dominican Republic are the sex ratios nearly equal. Brazil again shows evidence of transference of women into, rather than out of, the eligible age range.

In order to determine if the characteristics of women excluded from eligibility were different from those who were included, a comparison was made (not shown) of the percent rural among women in age groups 45-49 and 50-54 for those surveys in which the upper boundary effect was most pronounced. Since rural women are less likely to know their exact ages than urban women, one would expect that exclusion of women at the upper age boundary would be greater among rural than urban women. This would result in the percent rural being higher at age group 50-54 than at 45-49. In fact, large differences were observed for Burundi, Ghana, and Liberia, suggesting that interviews with rural women are particularly susceptible to the types of biases that result in the exclusion of older women from the eligible age range. On the other hand, the opposite was found in Mali, where eligible women in urban areas were more likely to be excluded.

Lower Boundary Effect. The lower limit of age eligibility in all surveys was 15 years. If the ages of young women were systematically understated in order to avoid eligibility, the age ratio for the 15-19 group would be low and the sex ratio high in comparison with the 10-14 group. The opposite would be true if women had erroneously been included in the survey.

[4] The age ratio is the number of women in the reference age group divided by the sum of the number of women in the two immediately adjacent age groups, multiplied by 100. The sex ratio is the number of men divided by the number of women in the reference age group, multiplied by 100.

[5] Since the upper boundary for age eligibility in Brazil and Guatemala was 45 years, age and sex ratios were calculated for age groups 40-44 and 45-49 (Table 2.1).

In only one survey, the Dominican Republic, is the age ratio for the 15-19 age group 110 or more, suggesting that some women under 15 years of age were included in the eligible age range. Age ratios for the 15-19 age group are lower than 90 in six countries (Botswana, Burundi, Ghana, Mali, Togo, and Peru), indicating that women may have been excluded from the eligible age range in these surveys. When the criterion of a ten-point drop in age ratios between age groups just outside (10-14) and just inside (15-19) the boundaries is considered (as was done in the World Fertility Survey evaluation), the list of surveys increases by two (Ecuador and Guatemala). Sex ratios of over 100 at age group 15-19 and of less than 100 at age group 10-14 further support the conclusion that misreporting of age has been selective for women in certain countries, most notably those in sub-Saharan Africa, particularly Burundi and Togo. None of the surveys in North Africa and Asia show strong evidence of either exclusion or inclusion at the lower boundary of age eligibility.

Boundary Effect Indices

In order to summarize the extent of distortion in the age/sex structures of the DHS household samples near the age eligibility boundaries, indices were calculated based on age and sex ratios. Three indices are considered: L, reflecting lower boundary distortion; U, reflecting upper boundary distortion; and T, which summarizes upper and lower boundary distortion. The L and U indices are defined as:

$$(AR_i - AR_o) - (SR_i - SR_o)$$

where AR and SR are the age ratios and sex ratios, subscript i denotes the age groups inside the boundary (i.e., 15-19 and 45-49), and subscript o denotes the age groups outside the boundary (i.e., 10-14 and 50-54). A positive sign indicates that too many women were considered eligible (in-transference) and a negative sign indicates that too many were considered ineligible (out-transference).

The index for total boundary distortion, T, is calculated by summing the values of L and U, disregarding the sign: ($T = |L| + |U|$). Because the movement of women at one boundary can be offset by movement at the other boundary, the T index indicates only the degree of distortion, not its direction. The values of these indices are presented in Table 2.2. In the table below the surveys are ranked based on the T index by the amount of distortion present.[6] Within categories, countries are listed in order of increasing values of the index.

[6] To facilitate comparison with the WFS surveys, the same categories are used here as in the WFS analysis (Rutstein, 1985).

Boundary Distortion Based on the T Index (T)

Negligible (T 0-24)	Low (T 25-49)	Moderate (T 50-99)	High (T 100+)
Indonesia	Morocco	Dominican Rep.	Sri Lanka
	Thailand	Zimbabwe	Guatemala
	Mexico	Brazil	Tunisia
	Colombia	Peru	Trinidad & Tob.
		Ghana	Senegal
			Mali
			Uganda
			Burundi
			Liberia
			Togo
			Ecuador
			Botswana

Comparison of DHS Data with Other Sources

So far, the assessment of age boundary effect has been entirely internal, comparing expected patterns—i.e., smooth declines in the size of age groups with increasing age—with observed patterns. One problem with this approach is that at the upper age boundary it is impossible to separate the effects of pushing women from age group 45-49 to 50-54 from the effects of heaping on age 50. To distinguish between the general tendency of a population to report ages ending in certain digits and the boundary effect that is more common with DHS and similar surveys, DHS household age data were compared with age data collected for the same populations from two other sources: WFS surveys and recent national censuses.

Table 2.2 gives the values of the indices: L, U, and T, for 22 DHS surveys, 13 WFS surveys, and 18 national censuses for which data were available.[7] For two countries, Mali and Uganda, neither WFS data nor recent census data were available.

Some patterns emerge from these comparisons. First, the pronounced upper boundary effect observed in most DHS surveys cannot be explained by general age misreporting patterns as reflected in the recent census figures of those countries. Since censuses do not involve conducting in-depth interviews with people in a particular age group, there is no incentive to push people from one age group to another. Thus, the census age data are presumably affected only by the problem of age heaping. The fact that in every country except Morocco, Brazil and Guatemala, the values of the U index derived from censuses are higher (closer to zero) than those from the DHS surveys is compelling evidence that the DHS data are subject to the boundary effect.

[7] The census indices are based on data available from the U.S. Bureau of the Census as of July 1990; only census data collected within 10 years of the DHS survey were considered.

Table 2.2 Indices of age eligibility distortion based on household data from DHS and WFS surveys and national censuses

Country	Lower Boundary (L)			Upper Boundary (U)			Total(T) (|L| + |U|)		
	DHS	WFS	Census	DHS	WFS	Census	DHS	WFS	Census
SUB-SAHARAN AFRICA									
Botswana	-67	--	- 2	-140	--	3	207	--	5
Burundi	-53	--	24	- 93	--	-22	146	--	47
Ghana	-27	- 4	a	- 67	-122	a	94	126	a
Liberia	- 5	--	33	-165	--	-10	170	--	43
Mali	-26	--	a	- 86	--	a	112	--	a
Senegal	- 1	32	a	-109	- 53	a	110	85	a
Togo	-59	--	42	-122	--	14	181	--	56
Uganda	1	--	a	-143	--	a	144	--	a
Zimbabwe	-16	--	1	- 42	--	6	58	--	7
NORTH AFRICA									
Morocco	11	19	11	- 19	-10	-50	30	29	60
Tunisia	4	4	14	-104	-73	9	108	77	23
ASIA									
Indonesia	- 3	29	18	-21	-99	-10	24	128	28
Sri Lanka	11	-21	0	-91	50	6	102	71	6
Thailand	3	11	5	-35	-13	- 3	38	24	8
LATIN AMERICA/CARIBBEAN									
Brazil	31	--	9	43	--	-10	74	--	19
Colombia	32	- 8	21	- 16	- 3	- 6	48	11	26
Dominican Republic	22	4	14	- 28	- 4	-10	50	8	24
Ecuador	-16	5	- 2	-171	-52	-14	187	57	16
Guatemala	-17	--	17	86	--	- 8	103	--	25
Mexico	- 9	- 5	5	- 32	-20	2	41	25	7
Peru	-38	5	2	- 48	- 6	3	86	11	6
Trinidad & Tobago	-10	- 5	18	-100	-75	- 8	110	80	26

[a]Census data within ten years of the DHS survey are not available.

For those countries with large amounts of upper age boundary exclusion in the DHS surveys (values of U less than -90) and for which census data were also available, (Botswana, Burundi, Liberia, Togo, Tunisia, Sri Lanka, Ecuador, and Trinidad and Tobago), the census value of U was never less than -22; five of these eight countries had census values that were within 10 points of the value zero, which indicates that no boundary distortion occurred in the census. Although the WFS data tend toward a DHS-type upper boundary effect, the effect is much smaller (with the exception of Ghana and Indonesia[8]).

Second, comparisons of DHS and census values for lower boundary distortion show similar results, namely that the values of L are generally much lower for the DHS surveys than for either the WFS surveys or censuses. In the DHS countries that show the greatest amount of lower boundary effect (values of L less than -35 in Botswana, Burundi, Togo, and Peru), the census L values are either near to or greater than zero. In countries for which a DHS/WFS comparison can be made, no clear conclusions can be made except that, on the whole, lower boundary effects are considerably less pronounced than upper boundary effects in both DHS and WFS surveys.

In sum, the results of this comparative assessment of the quality of DHS data regarding age eligibility boundaries indicates that in several surveys out-transference of women 45-49 from eligibility has been greater in DHS surveys than in other data sources. The evidence for lower boundary exclusion is much less compelling, although age patterns in a few countries also indicate out-transference from the 15-19 age group.

2.3 REPORTING OF HOUSEHOLD RESIDENCY

In most DHS surveys, a second criterion for eligibility is that the woman slept in the selected dwelling unit during the night preceding the household interview (de facto sample). In the household interview, all persons considered

[8] One reason that the survey in Indonesia shows so little evidence of exclusion may be that interviewers were paid a fixed amount for each household and individual questionnaire they completed. If anything, this practice might have encouraged interviewers to increase rather than decrease the number of eligible women.

residents, as well as visitors and guests, are listed on the household questionnaire. For each person listed, the interviewer ascertains whether he/she slept in the dwelling during the previous night.[9] The de facto approach is used to affix a probability of selection to all women and to avoid the possibility that women may be selected twice. In theory, most residents who slept away had a chance to be interviewed in other households, as eligible visitors "sleeping in the household." (The small number of women residents sleeping in hotels, hospitals, etc. would not have had a chance to be interviewed.)

If large differences exist between the number of residents sleeping away and the number of overnight visitors (nonresidents, but eligible), there may be a problem of systematic exclusion of women who are "sleeping away" (from their resident household).[10] Given the significant rigors of fieldwork in developing countries, interviewers may intentionally report that some resident women—who were away at work, at market, or simply at a neighbor's house—did not sleep in the household the preceding night in order to avoid having to revisit the household for another interview (as required by DHS protocol). When this type of systematic exclusion takes place, the result is a sample of women that excludes a disproportionate number of women who are more frequently away from the home.

Table 2.3 shows by age group (A) the number of women reported to reside in the households for which an interview was conducted, (B) the percentage of resident women who were reported to have slept away from their dwelling the previous night, and (C) the ratio of overnight visitors to

[9] In Peru, only persons who slept in the dwelling the night before were placed on the original household list, thus excluding from the survey those residents who slept away that night.

[10] Small differences are expected because some women are away at boarding school, in the hospital, in jail, or at locations otherwise excluded from the sampling frame. However, the proportion of women in these categories is small in developing countries.

Table 2.3 (A) Number of women reported to reside in interviewed households, (B) the percentage of resident women not sleeping in the household during the night before the survey, and (C) the standardized percentage of nonresident women sleeping in the household during the night before the survey, by five-year age group, Demographic and Health Surveys, 1986-1989

		Age Group								
Country		10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54
SUB-SAHARAN AFRICA										
Botswana	A	1837	1085	1031	960	710	522	358	282	425
	B	4.2	14.0	12.6	11.9	10.0	13.0	11.2	13.5	7.8
	C	2.6	4.4	4.9	5.0	3.4	5.4	0.8	3.9	4.2
Burundi	A	1294	894	882	879	654	500	290	297	404
	B	4.3	17.8	12.1	7.2	6.4	4.0	5.5	4.7	1.0
	C	1.7	2.7	3.1	3.1	1.5	2.0	1.4	0.6	0.7
Ghana	A	1405	970	970	952	708	585	401	400	409
	B	3.2	13.1	12.2	11.4	9.6	10.4	8.2	10.0	6.6
	C	0.9	2.7	4.4	4.2	2.5	2.2	1.7	0.8	2.0
Liberia	A	1635	1253	1105	1180	746	713	390	421	696
	B	5.0	12.6	12.0	11.9	15.0	14.9	16.9	15.0	12.8
	C	1.3	4.5	5.7	5.0	3.4	2.4	2.6	3.6	3.2
Mali	A	994	615	584	665	530	446	329	277	325
	B	4.3	17.4	9.8	7.7	5.8	7.8	7.0	5.4	4.9
	C	0.5	1.6	0.7	1.1	0.8	0.9	1.2	0.0	0.3
Senegal	A	1879	1510	1341	1221	931	744	478	453	639
	B	5.2	10.1	10.5	10.5	7.6	10.8	9.6	14.8	5.5
	C	2.0	3.2	5.1	4.8	3.9	3.5	3.6	2.9	4.1
Togo	A	1272	765	663	631	504	397	288	254	411
	B	3.2	8.9	5.1	5.7	5.8	7.3	6.9	5.5	4.4
	C	2.0	4.7	6.3	5.4	2.2	1.3	1.7	1.6	2.7
Uganda	A	1688	1258	1061	922	664	467	325	278	389
	B	5.2	10.4	9.1	8.6	8.3	7.7	10.8	9.4	8.5
	C	1.6	6.4	4.5	4.6	2.7	2.1	4.0	4.0	1.3
Zimbabwe	A	1582	1157	880	751	634	497	340	312	338
	B	3.9	10.5	9.7	10.8	9.5	8.5	11.8	9.0	5.0
	C	2.1	7.8	12.7	9.6	5.5	4.0	3.3	4.8	5.9

Table 2.3 (continued)

Country		Age Group 10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54
NORTH AFRICA										
Morocco	A	2723	2466	1972	1793	1361	1024	765	813	840
	B	3.5	5.1	6.9	8.5	6.8	5.5	7.2	6.3	10.2
	C	2.5	3.4	5.0	4.7	2.9	2.3	1.8	2.3	3.2
Tunisia	A	1861	1683	1573	1231	1052	822	611	452	643
	B	1.6	2.3	3.1	2.7	1.9	1.7	2.3	1.8	3.0
	C	2.8	3.5	4.7	4.0	4.8	1.7	3.8	1.1	1.6
ASIA										
Indonesia[a]	A	4082	3599	3125	2772	2205	1732	1456	1498	1455
Sri Lanka	A	2202	2132	1989	1661	1493	1357	1021	727	899
	B	1.6	4.2	7.1	4.6	5.1	3.8	2.6	4.1	3.4
	C	0.8	1.8	2.0	2.8	1.7	1.5	0.7	0.7	1.2
Thailand	A	2464	2477	2258	1784	1739	1391	1083	910	922
	B	4.0	9.3	8.5	6.8	5.2	4.4	5.6	3.8	5.3
	C	2.4	3.4	3.5	3.1	2.4	1.4	1.8	2.6	1.8
LATIN AMERICA/CARIBBEAN										
Brazil[a]	A	1545	1551	1381	1163	1037	879	714	444	328
Colombia	A	1247	1295	1191	963	761	674	452	417	397
	B	2.0	3.5	3.1	3.7	2.1	3.7	3.3	2.6	2.8
	C	1.4	2.0	2.4	1.9	1.6	1.0	0.9	0.7	0.8
Dominican Republic	A	2234	2254	1896	1394	1094	871	671	605	610
	B	3.3	6.0	6.1	4.7	4.3	3.8	5.4	4.3	4.4
	C	2.1	4.1	4.4	3.3	2.9	2.1	2.8	2.1	1.6
Ecuador	A	1468	1149	999	910	718	630	439	328	560
	B	1.7	6.1	6.0	5.1	4.7	7.0	5.7	4.9	3.6
	C	1.2	2.0	2.6	2.0	1.1	1.4	0.9	0.3	0.5
Guatemala	A	1253	916	739	714	571	511	347	429	326
	B	4.0	9.0	5.1	3.8	3.0	4.3	3.5	3.5	2.1
	C	0.4	1.2	1.1	0.8	0.5	0.2	0.3	0.7	0.3
Mexico	A	2857	2389	1813	1697	1340	1110	899	710	688
	B	1.6	6.9	6.1	3.3	2.8	1.9	2.1	3.5	2.6
	C	0.8	2.8	3.4	1.4	0.9	1.3	0.2	0.0	1.0
Peru[b]	A	1599	1162	1019	853	718	612	509	409	436
Trinidad & Tobago	A	871	798	868	851	624	494	414	315	410
	B	3.7	7.1	7.8	5.2	5.6	5.1	3.6	3.5	2.9
	C	3.1	3.0	3.0	2.5	1.1	0.8	1.0	0.6	0.7

[a]De jure sample
[b]Sample does not include resident women who did not sleep in the household during the night before the survey.

residents, expressed as a percentage.[11] The values for B and C should be roughly equal; however, if large numbers of women are excluded—falsely identified as "sleeping away,"—B will be much larger than C. The "net exclusion," defined as the difference between B and C, is an estimate of the level of bias associated with this phenomenon, and represents women who are not reported as sleeping in any household (and, therefore, did not have a chance to be interviewed). The estimated net exclusions are given in Table 2.4 and shown graphically in Figure 2.3 (the shaded areas).

[11] In two surveys, Brazil and Indonesia, a de jure sample was used, i.e., only residents were eligible. In these surveys, both residents who slept in the dwelling the previous night and residents who did not, were to be interviewed with the individual questionnaire. Although, this procedure also ensures that all women have an equal (single) chance of being selected, in actual practice, interviewing residents who did not sleep in the dwelling the previous night was often impossible, since it would require following them to their current location. In any case, the rationale used in this analysis does not apply to these two surveys.

Table 2.4 Estimated percentage of women who were excluded from eligibility for the individual interview due to misrecording of "sleeping away" status, by age group, Demographic and Health Surveys, 1986-1989

Country	Age Group								
	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54
SUB-SAHARAN AFRICA									
Botswana	1.6	9.6	7.7	6.9	6.6	7.6	10.4	9.6	3.6
Burundi	2.6	15.1	9.0	4.1	4.9	2.0	4.1	4.1	0.3
Ghana	2.3	10.4	7.8	7.2	7.1	8.2	6.5	9.2	4.6
Liberia	3.7	8.1	6.3	6.9	11.6	12.5	14.3	11.4	9.6
Mali	3.8	15.8	9.1	6.6	5.0	6.9	5.8	5.4	4.6
Senegal	3.2	6.9	5.4	5.7	3.7	7.3	6.0	11.9	1.4
Togo	1.2	4.2	-1.2	0.3	3.6	6.0	5.2	3.9	1.7
Uganda	3.6	4.0	4.6	4.0	5.6	5.6	6.8	5.4	7.2
Zimbabwe	1.8	2.7	-3.0	1.2	4.0	4.5	8.5	4.2	-0.9
NORTH AFRICA									
Morocco	1.0	1.7	1.9	3.8	3.9	3.2	5.4	4.0	7.0
Tunisia	-1.2	-1.2	-1.6	-1.3	-2.9	0.0	-1.5	0.7	1.4
ASIA									
Indonesia[a]	--	--	--	--	--	--	--	--	--
Sri Lanka	0.8	2.4	4.9	1.8	3.4	2.3	1.9	4.4	2.2
Thailand	1.6	5.9	5.0	3.7	2.8	3.0	3.8	1.2	3.5
LATIN AMERICA/CARIBBEAN									
Brazil[a]	--	--	--	--	--	--	--	--	--
Colombia	0.6	1.5	0.7	1.8	0.5	2.7	2.4	1.9	2.0
Dominican Republic	1.2	1.9	1.7	1.4	1.4	1.7	2.6	2.2	2.8
Ecuador	0.5	4.1	3.4	3.1	3.6	5.6	4.8	4.6	3.1
Guatemala	3.6	7.8	4.0	3.0	2.5	4.1	3.2	2.8	1.8
Mexico	0.8	4.1	2.7	1.9	1.9	0.6	1.9	3.5	1.6
Peru[b]	--	--	--	--	--	--	--	--	--
Trinidad & Tobago	0.6	4.1	4.8	2.7	4.5	4.3	2.6	2.9	2.2

Note: The values in this table are calculated as the difference between B and C values presented in Table 2.3.

[a]De jure sample
[b]Residents only interviewed, "sleeping-away" status not ascertained

The results of this analysis strongly suggest that eligible women are being excluded from the opportunity to be interviewed; this was found to be especially true in surveys in sub-Saharan Africa. The net exclusion is as high as 15 percent for women 15-19 in Burundi and Mali, and is 5 percent or over for all eligible age groups in Botswana, Ghana, Liberia, and Mali. In other regions, only the surveys in Thailand, Ecuador, Guatemala, and Trinidad and Tobago indicate substantial misrecording of "sleeping-away" status.

Rather than intentionally overstate or understate the respondent's age in order to "push" the woman over the age eligibility boundary, some interviewers may have opted to record that the woman did not sleep in the household the previous night. This possibility is examined in Table 2.5 which shows the difference in net exclusion between age groups that straddle the age eligibility boundaries. If deliberate misclassification of where women spent the preceding night has occurred (in order to reduce the interviewing load), it would be expected that net exclusion would be larger for the age groups immediately inside the age eligibility boundaries (15-19 and 45-49) than for the age groups immediately outside the boundaries (10-14 and 50-54).

Of the 19 de facto surveys which have data on women's presence in the household the previous night, 13 show more than a one percent difference in the net percentage excluded at the lower age boundary. While some difference in exclusions at the lower boundary can be expected—some of the young women this age may be away at boarding school but are still considered residents of the household—differences of more than five percent in Botswana, Burundi, Ghana, and Mali, appear excessive and suggest deliberate exclusions.

At the upper boundary, 10 surveys show differences in net exclusion exceeding one percent and three show differences greater than five percent (Botswana, Senegal, and Zimbabwe). Unlike at the lower boundary, major changes in life

Figure 2.3 Percentage of women not reported as sleeping in a household during the night before the survey by age, Demographic and Health Surveys, 1986-1989

SUB-SAHARAN AFRICA

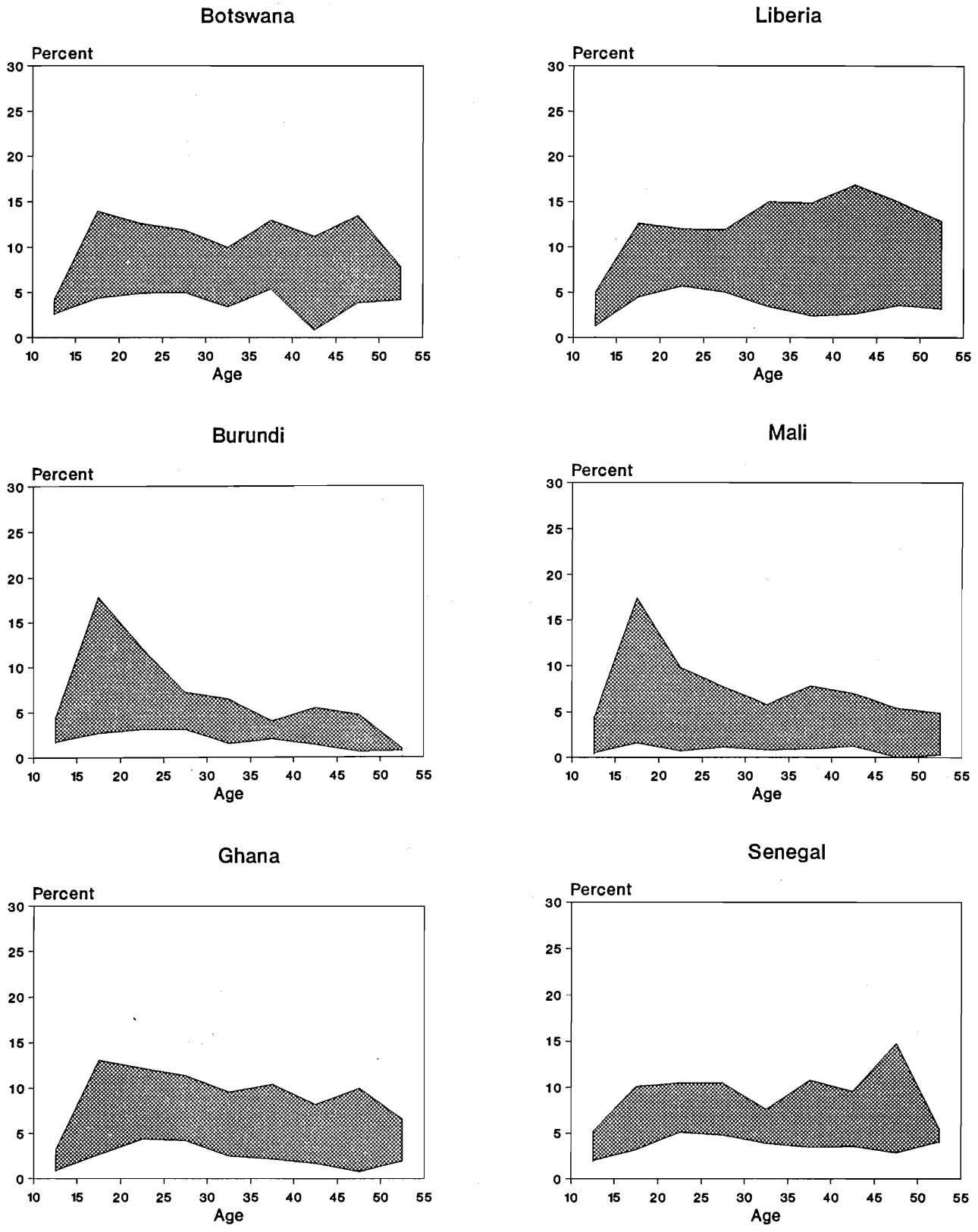
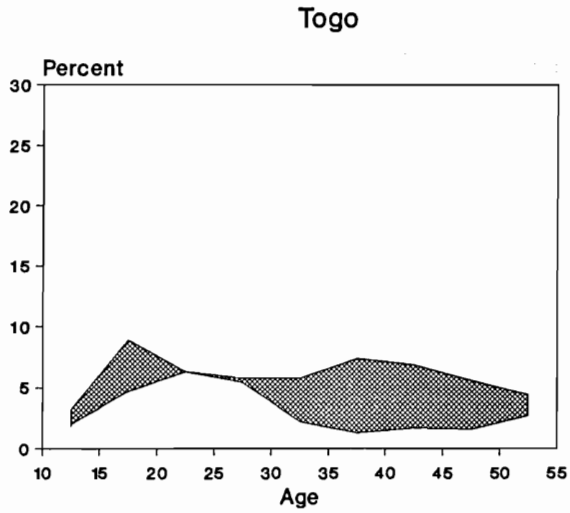
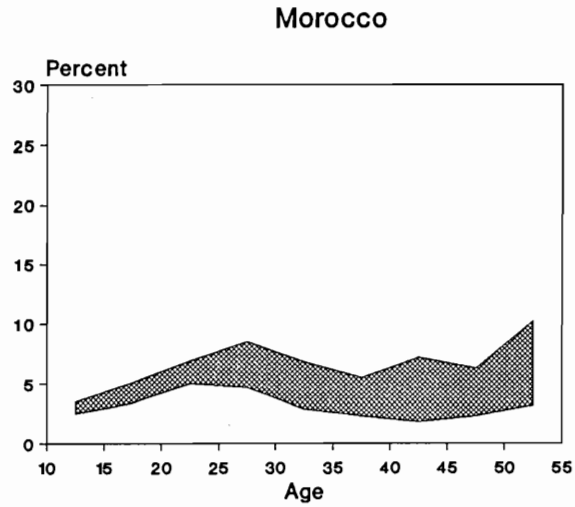
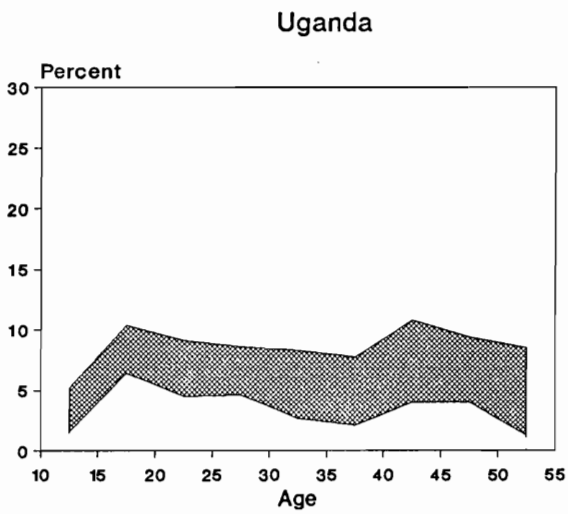


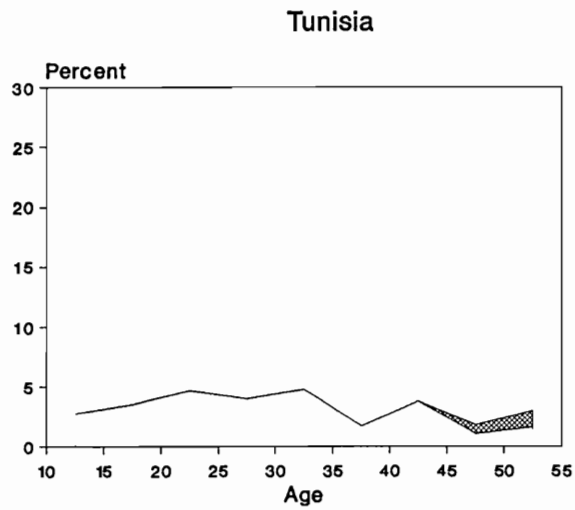
Figure 2.3—*Continued*

NORTH AFRICA









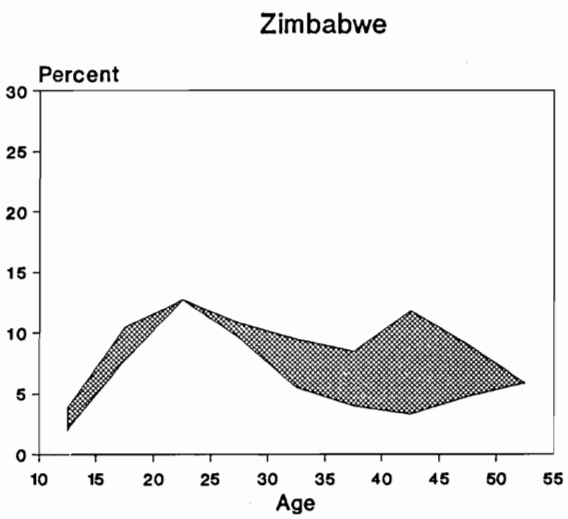
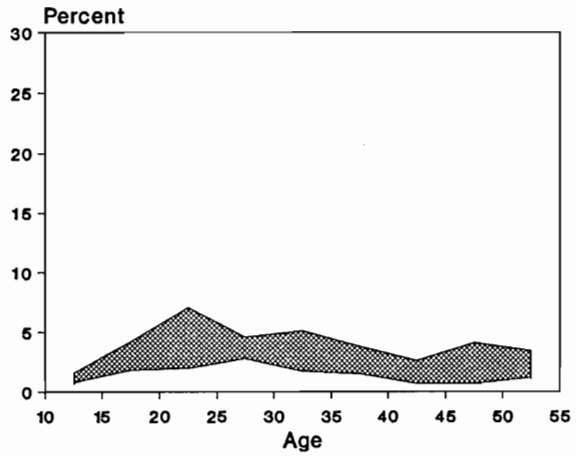


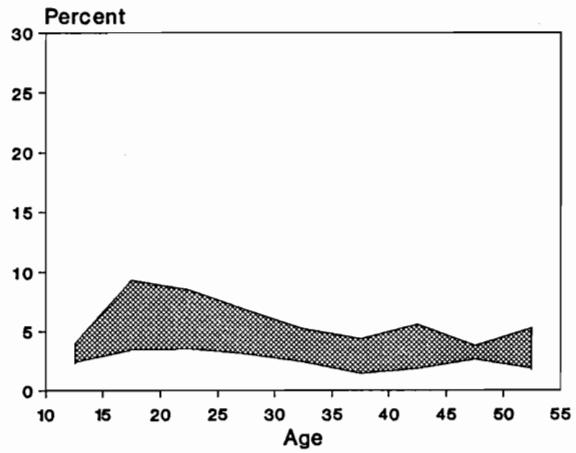
Figure 2.3—*Continued*

ASIA

Sri Lanka

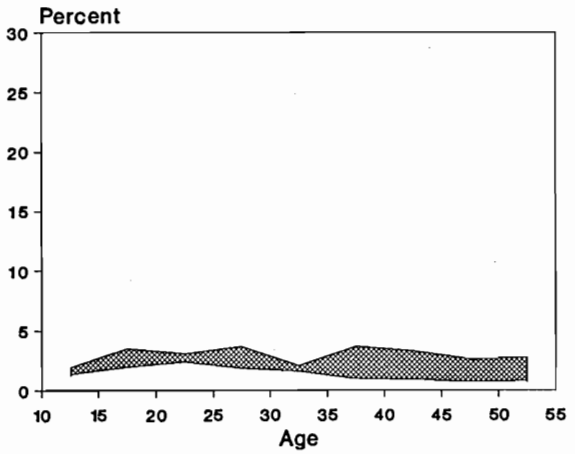


Thailand

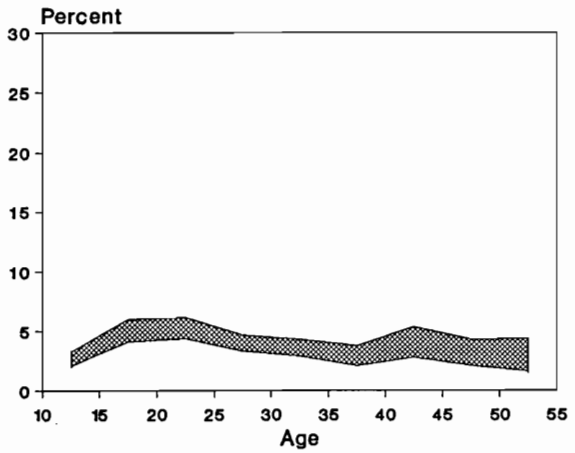


LATIN AMERICA/CARIBBEAN

Colombia



Dominican Republic



Ecuador

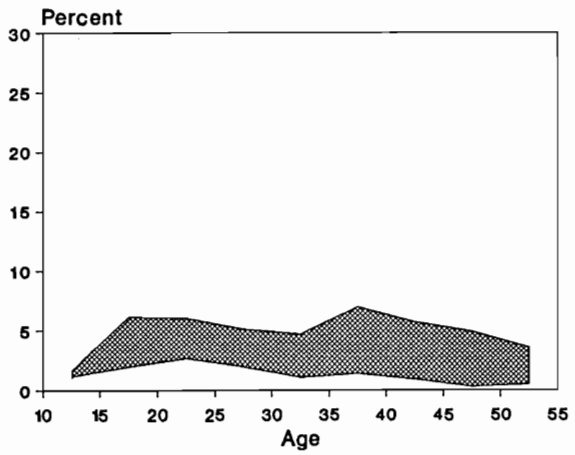
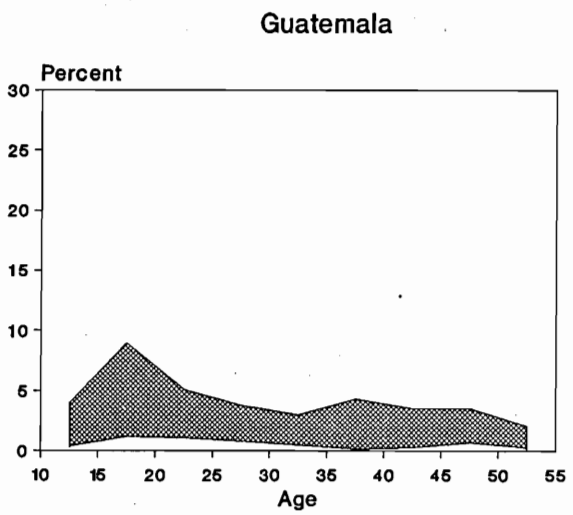
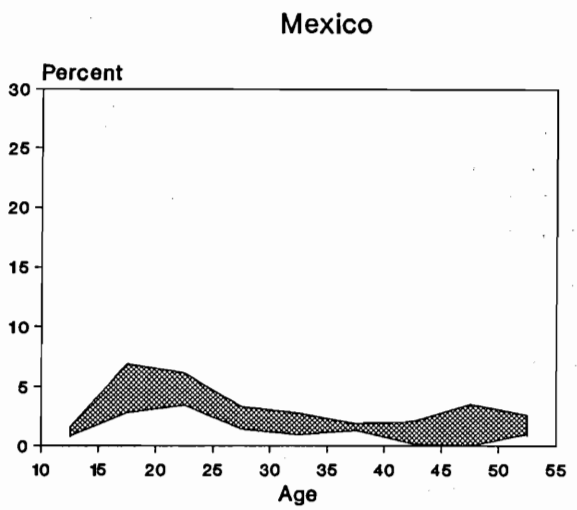


Figure 2.3—*Continued*





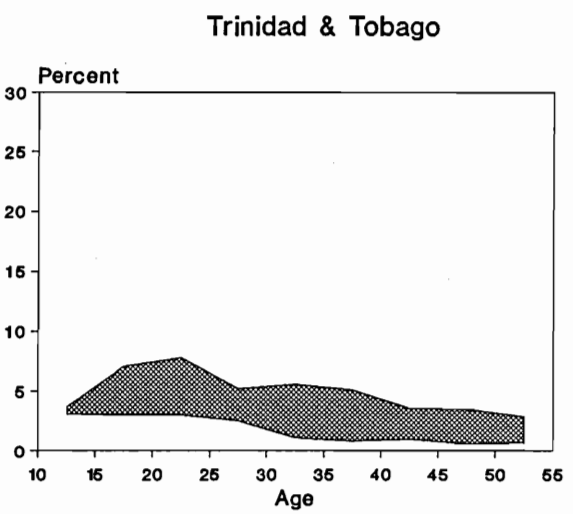


Table 2.5 Difference in the estimated percentage of women who were excluded from eligibility due to misrecording of "sleeping away" status, between age groups immediately inside and outside the age eligibility boundaries, Demographic and Health Surveys, 1986-1989

Country	[15-19] - [10-14]	[45-49] - [50-54]
SUB-SAHARAN AFRICA		
Botswana	8.0	6.0
Burundi	12.5	3.8
Ghana	7.1	4.6
Liberia	4.7	1.8
Mali	12.0	0.8
Senegal	3.7	10.5
Togo	3.0	2.2
Uganda	0.4	-1.8
Zimbabwe	0.9	5.1
NORTH AFRICA		
Morocco	0.7	-3.0
Tunisia	0.0	-0.7
ASIA		
Indonesia[a]	--	--
Sri Lanka	1.6	1.2
Thailand	4.3	-2.3
LATIN AMERICA/CARIBBEAN		
Brazil[a]	--	--
Colombia	0.9	-0.1
Dominican Rep.	0.7	-0.6
Ecuador	3.6	1.5
Guatemala	4.2	0.4[b]
Mexico	3.3	1.9
Peru[c]	--	--
Trinidad & Tobago	3.5	0.7

[a]De jure sample
[b]Age groups 40-44 and 45-49
[c]"Sleeping away" status not ascertained

style between age groups at the upper age boundary (in this case 45-49 and 50-54), would not be expected. Figure 2.3 indicates that the reductions in net exclusion (shaded areas) from age group 45-49 to age group 50-54 is due solely to the drop in the percentage of women reported as having slept away from their households, and not to an increase in visitors. This suggests intentional exclusion of women by interviewers.

In order to classify surveys according to the estimated overall level of exclusion due to misreporting of sleeping-away status of eligible women, net exclusion was calculated for all women 15-49. (Calculations for Guatemala are based on women 15-44.) In the table below the surveys are ranked according to net exclusion due to "sleeping away."

Estimated net exclusion of eligible women due to misrecording of "Sleeping-away" status

Negligible (<2.5%)	Low (2.5-4.9%)	Moderate (5.0-7.4%)	High (7.5%+)
Tunisia (-1.3)	Mexico (2.6)	Senegal (6.3)	Botswana (8.1)
Colombia (1.5)	Togo (2.7)	Burundi (7.2)	Ghana (8.3)
Dominican Republic (1.8)	Sri Lanka (3.0)		Mali (8.3)
Zimbabwe (2.3)	Morocco (3.0)		Liberia (9.2)
	Trinidad & Tobago (3.8)		
	Thailand (4.0)		
	Ecuador (4.0)		
	Guatemala (4.3)		
	Uganda (4.8)		

In six of the 19 surveys (all in sub-Saharan Africa), more than five percent of the women are estimated to have been excluded because the interviewer falsely reported them as not sleeping in the household during the night before the interview. The results for Tunisia are unusual in that more overnight visitors were recorded than women sleeping away from home. A possible explanation for this phenomenon is that many Tunisians residing outside the country (particularly in Europe) return in the summer for visits. Since the DHS survey was carried out in the summer, the negative percentages in Table 2.4 may reflect the presence of vacationing "visitors."

2.4 SIMULATING THE EFFECTS OF EXCLUSION

The DHS surveys are designed to provide nationally representative estimates of important demographic and health parameters. Any systematic exclusion of women from the individual interview raises questions of selection bias. Two factors determine whether the final estimates are seriously biased: (1) the level (amount) and age pattern of exclusion, and (2) the characteristics of excluded women in relation to those included, with respect to the parameters being estimated.

The following simulation is designed to assess the potential effects of bias due to exclusion on the estimation of (1) the total fertility rate in the last five years, (2) under-five mortality in the last five years, and (3) current contraceptive prevalence among women in union. The premise of the exercise is that departures from expected patterns in the data do not reflect real characteristics of the population, but rather, are caused wholly by exclusion. It has already been shown that this premise is sound in the case of age eligibility boundary effects (see section 2.2). The simulation considers the three aspects of exclusion described above; namely, (1) lower age boundary effects, (2) upper age boundary effects, and (3) the effects of misclassification of "sleeping-away" status.

The first step in the simulation was to adjust the DHS samples in order to "recover" the excluded women. In the lower boundary and upper boundary adjustments, the age distributions were effectively smoothed at the age eligibility boundaries, resulting in most cases in the addition of women to the age groups immediately inside the boundaries. The number of women excluded due to "sleeping away" (by age group) was determined and added to the sample based on the net exclusion figures already presented in Table 2.4.

The second step in the simulation was to assign to the women (and their children, in the case of childhood mortality) age-specific probabilities related to the three demographic parameters being examined. The specific assumptions are given below.

ASSUMPTIONS:

Total Fertility Rate

Lower Boundary

FLB0 Excluded women had an age-specific fertility of 0.0 (no births)

Upper Boundary

FUB0 Excluded women had an age-specific fertility of 0.0

FUB2 Excluded women had twice the age-specific fertility as included women

Sleeping Away

FSA75 Excluded women had 75 percent of the age-specific fertility of included women

FSA125 Excluded women had 125 percent of the age-specific fertility of included women

Under-five Mortality ($_5q_0$) - Under same age-specific fertility as included women.

Lower Boundary

MLB150 Excluded children have 150 percent the rate of under-five mortality by age of mother as included children

Upper Boundary

MUB150 Excluded children have 150 percent the rate of under-five mortality by age of mother as included children

Sleeping Away

MSA150 Excluded children have 150 percent the rate of under-five mortality by age of mother as included children

Contraceptive Use - Under same age-specific union status as included women.

Lower Boundary

CLB1 Same contraceptive use rates as included women

Upper Boundary

CUB1 Same contraceptive use rates as included women

Sleeping Away

CSA1 Same contraceptive use rates as included women

Results of the Simulations

The results of the simulations are presented in Table 2.6 as percentage changes in the parameters resulting from the application of these assumptions and procedures. A negative value in the table indicates that the simulated rate was lower by that percentage than the observed rate.

Looking first at change in the total fertility rate (TFR), it appears that only under the most extreme set of assumptions is appreciable change (bias) evident. In the scenario whereby women excluded on the basis of age are assumed to have no fertility and women excluded on the basis of "sleeping-away" status have age-specific fertility levels that are 75 percent of those of interviewed women (combined FLB0, FUB0, and FSA75), the simulated estimates of the TFRs are less than 5 percent lower than the observed rates in all cases except Botswana (-5.1 percent). Countries with the most change are those in sub-Saharan Africa; outside this region, the simulated rates are less than 2 percent lower than the observed rates in all countries except for Ecuador. The more pronounced changes observed in sub-Saharan Africa countries were due to the higher levels of exclusion based on "sleeping-away" status. Age boundary exclusion does not substantially bias the TFR estimates because the fertility of women in the boundary age groups (15-19 and 45-49) is normally low.

Simulation of the effects of exclusion on estimates of under-five mortality produced a pattern of results across countries similar to that of the TFR simulation, except that the changes occur in the opposite direction: the effect of exclusion is typically to bias mortality downward and fertility upward under the range of assumptions employed here. Again, only in the most extreme scenario of combined age and "sleeping-away" exclusions are the results noteworthy, and even then only in some sub-Saharan countries (especially Mali, Ghana, and Liberia). For Togo, Uganda, Zimbabwe, and the countries outside sub-Saharan Africa, the maximum bias associated with exclusion was 2.5 percent (i.e., the simulated rates were 2.5 percent higher than the observed rates).

Under the assumptions used to simulate bias in the estimates of contraceptive prevalence, no appreciable changes were observed. The greatest amount of change occurred for Botswana where the contraceptive prevalence rate simulated under the combined boundary and "sleeping-away" status assumptions is 2 percent lower than the observed rate. It should be noted, however, that the changes are only due to the effect of changes in age composition, since excluded women were assumed to have the same (current) age-specific level of contraceptive use as women who were interviewed.

Response to the Individual Questionnaire

Overall, response rates for the individual questionnaire were high (see Table 2.7). In 12 of the 22 surveys, over 95 percent of the women declared eligible for the individual questionnaire were interviewed (more than 98 percent in six surveys). In nine surveys, the response rate was between 90 and 95 percent. Only in Brazil was the response rate for the individual questionnaire below 90 percent (88 percent). Given these high levels of response to the individual questionnaire, it is unlikely that bias due to individual nonresponse would have a substantial effect on estimates produced from the DHS data (with the possible exception of Brazil).

Table 2.6 Results of simulations to estimate the effect of lower boundary, upper boundary, and "sleeping away" exclusions on total fertility rate, under-five mortality, and contraceptive prevalence, Demographic and Health Surveys, 1986-1989

	Total Fertility Rate (Last Five Years)									Under-five Mortality Rate (Last Five Years)				Contraceptive Prevalence Rate[a]			
Country	FLB0	FUB0	FUB2	FSA75	FSA125	FLB0 FUB0 FSA75	FLB0 FUB0 FSA125	FLB0 FUB2 FSA75	FLB0 FUB2 FSA125	MLB150	MUB150	MSA150	MLB150 MUB150 MSA150	CLB1	CUB1	FSA1	FLB1 FUB1 FSA1
SUB-SAHARAN AFRICA																	
Botswana	-2.6	-1.0	+1.0	-2.0	+2.0	-5.1	-1.4	-3.4	+.4	+2.1	+.0	+3.6	+3.8	-.4	-1.5	-.3	-2.0
Burundi	-.4	-1.1	+1.1	-1.3	+1.3	-2.6	-.1	-.6	+1.9	+.6	+.2	+2.9	+3.5	-.1	+.5	-.5	+.0
Ghana	-.9	-.5	+.5	-1.9	+1.9	-3.1	+.6	-2.3	+1.4	+1.0	+.3	+3.7	+4.3	-.3	-.2	-.2	-.7
Liberia	-.4	-1.3	+1.3	-2.2	+2.2	-3.8	+.5	-1.3	+2.9	+.4	+.3	+3.4	+4.1	-.2	+1.1	+.2	+1.0
Mali	-1.8	-.0	+.0	-1.9	+1.9	-3.8	-.1	-3.0	+.7	+1.7	+.1	+4.3	+4.9	+.9	-1.3	+.8	+.4
Senegal	-.1	-.9	+.9	-1.9	+1.9	-2.7	+1.1	-1.2	+2.6	+.1	+.2	+3.6	+4.0	-.1	+.0	-.5	-.5
Togo	-1.4	-1.5	+1.5	-.6	+.6	-3.3	-2.1	-.5	+.7	+1.4	+.3	+.6	+1.5	-.2	-.2	+.5	+.1
Uganda	-.4	-.5	+.5	-1.2	+1.2	-2.0	+.3	-1.2	+1.2	+.6	+.1	+2.1	+2.4	-.2	-.0	+.3	+.0
Zimbabwe	-.2	-.2	+.2	-.5	+.5	-1.0	+.0	-.5	+.5	+.3	+.1	+1.0	+1.3	-.0	-.4	-.5	-.9
NORTH AFRICA																	
Morocco	+.0	+.0	+.1	-.8	+.8	-.8	+.8	-.7	+.9	-.1	-.0	+1.5	+1.5	+.0	-.1	+.1	+.0
Tunisia	+.0	-.3	+.3	+.4	-.4	+.1	-.7	+.7	-.1	-.0	+.2	-1.0	-.6	-.0	-.1	+.0	-.1
ASIA																	
Indonesia[b]	-.3	+.0	-.0	--	--	--	--	--	--	+.3	-.0	--	--	-.1	+.2	--	--
Sri Lanka	+.1	-.1	+.1	-.8	+.8	-.8	+.8	-.6	+1.0	-.1	+.0	+1.6	+1.6	+.0	-.2	-.1	-.3
Thailand	+.0	-.1	+.1	-1.1	+1.1	-1.2	+.9	1.0	+1.1	+.0	+.1	+2.0	+2.2	-.0	-.2	-.1	-.3
LATIN AMERICA/CARIBBEAN																	
Brazil[b]	+.6	+.4	-.4	--	--	--	--	--	--	-.5	-.3	--	--	+.0	+.1	--	--
Colombia	+.8	.0	.0	-.4	+.4	+.4	+1.1	+.4	+1.2	-1.3	+.0	+.7	+.5	+.2	-.1	+.0	+.1
Dominican Republic	+.6	-.1	+.1	-.5	+.5	+.0	+1.0	+.2	+1.1	-.8	+.0	+.9	+.8	+.2	-.1	+.0	+.0
Ecuador	-.6	-.9	+.9	-1.0	+1.0	-2.4	-.4	-.7	+1.2	+.3	+.4	+1.6	+2.2	-.2	-1.2	+.0	-1.4
Guatemala	-.3	-.9	+.9	-.7	+.7	-1.9	-.6	-.1	+1.3	+.4	+.8	+1.2	+2.5	-.2	+.4	-.2	-.1
Mexico	-.3	-.1	+.1	-.5	+.5	-.9	+.2	-.7	+.4	+.2	+.0	+1.2	+1.2	-.1	-.2	-.2	-.5
Peru[c]	-.9	-.2	+.2	--	--	--	--	--	--	+1.0	+.1	--	--	-.2	-.6	--	--
Trinidad & Tobago	-.4	-.1	+.1	-1.1	+1.1	-1.5	+.6	-1.3	+.8	+.5	+.0	+2.1	+2.3	-.0	-.4	+.2	-.2

Note: The results of the simulations are expressed as percent of change (bias) due to the adoption of selected assumptions concerning the excluded women (see pages 21-22).

[a]Current use of modern methods (all methods for sub-Saharan Africa) among women in union

[b]De jure sample

[c]Residents who did not sleep in the household during the night before the survey were not included in the household questionnaire.

Table 2.7 Percent distribution of eligible women by response to the individual interview, Demographic and Health Surveys, 1986-1989

Country	Percentage of Women Who Responded	Percentage of Women Who Refused	Other Nonresponse	Total	No. of Eligible Women
SUB-SAHARAN AFRICA					
Botswana	94.0	0.2	5.8	100.0	4648
Burundi	98.1	0.1	1.8	100.0	4077
Ghana	98.1	0.3	1.6	100.0	4574
Liberia	97.4	0.2	2.4	100.0	5340
Mali	98.6	0.1	1.3	100.0	3246
Senegal	96.5	0.0	3.5	100.0	4574
Togo	98.6	0.3	1.1	100.0	3409
Uganda	97.4	0.2	2.4	100.0	4730
Zimbabwe	94.0	0.1	5.9	100.0	4467
NORTH AFRICA					
Morocco	98.9	0.2	0.9	100.0	6050
Tunisia	96.7	0.2	3.1	100.0	4325
ASIA					
Indonesia	98.5	0.1	1.4	100.0	12065
Sri Lanka	95.1	0.2	4.7	100.0	6170
Thailand	94.1	NA	NA	--	7201
LATIN AMERICA/CARIBBEAN					
Brazil	87.5	2.5	10.0	100.0	6733
Colombia	94.2	NA	NA	--	5329
Dominican Republic	93.4	0.6	6.0	100.0	8186
Ecuador	94.9	1.7	3.4	100.0	4967
Guatemala	93.3	1.9	4.8	100.0	5532
Mexico	96.0	1.1	2.9	100.0	9709
Peru	94.6	NA	NA	--	5282
Trinidad & Tobago	90.7	2.0	7.3	100.0	4196

NA = Not available

3 Age Reporting in the Individual Questionnaire

Respondent's age in the individual questionnaire is obtained by asking for both birth date (year and month) and age in completed years.[12] The interviewer was instructed to "make a serious effort" to reconcile the two answers if inconsistent. If the respondent could not give either her birth date or her age, the interviewer was instructed to request a document indicating the respondent's age or to probe by estimating the respondent's age in relation to the ages of other members of the household, or in relation to the date of (or age at) her first marriage or first birth. If all else failed, the interviewer was instructed to guess the respondent's age.

In societies where vital registration is uncommon, and where ages and dates are not required in daily life, individuals often lack documents indicating age, and may not have a clear idea of their own ages—only when dealing with official matters may a report of age be required. For a woman, this situation is likely to be compounded by a low level of education and the fact that in many societies a woman is represented in official matters by her husband or father.

Even when a document is available to determine current age, unless it was obtained in early childhood, it may also lead to an erroneous report of age. This is because many documents obtained later in life are themselves based on crude estimates of age.

Lacking documentation or a response, the survey interviewer is forced to estimate current age on the basis of physical appearance or the milestones reached in a person's life. Unfortunately, appearances and milestones are usually themselves related to the phenomena under study and thus may bias the results. For example, in many cases, the respondent's current age is estimated based on the number of children she has borne, or the age of the oldest or youngest of her children. Estimating age on the basis of physical appearance is subjective and may be influenced by the interviewer's background. Typically, interviewers are young adults many of whom come from an urban setting. To them, a poor or rural woman may appear older than she actually is because of poor diet, lack of teeth, sun-wrinkled skin, or dirty appearance. Similarly a woman who has had many children may appear older than she is because of the strain of repeated pregnancy, breastfeeding, and child-rearing.

[12] The questions were "In what month and year were you born?" and "How old were you at your last birthday?"

When the age of an eligible woman is not known, interviewers are likely to place respondents at about the middle of the age range, usually between ages 30 and 40. A more detailed discussion of the causes and effects of age misreporting is presented by Ewbank (1981).

3.1 REPORTED AGE VS. IMPUTED AGE

Table 3.1 shows the percent distribution of women according to whether information on date of birth was given (reported or estimated) or imputed. The highest level of reported birth dates is for Latin America and the Caribbean, where over 95 percent of the respondents were able to give the year and month of their date of birth. At the other extreme, in eight surveys (six in sub-Saharan Africa), less than half of the women were able to report their date of birth. In Mali and Morocco, only 9 and 13 percent of women, respectively, could provide year and month of birth information. Unfortunately, there is no record of how age was determined in the cases where the birth date was not obtained.

3.2 DIGIT PREFERENCE

In the reporting of age, a tendency to prefer certain digits, commonly zero and five, is indicative of incorrect ascertainment of age. The greater the amount of this "age heaping," the lower the confidence in the quality of the age data. The Myers index is used here to measure the extent of age heaping in the DHS surveys. The age range over which the index is calculated is restricted to 20-49 years so that each digit will have approximately the same chance (i.e., 10 percent) of occurring in a true distribution. No "blending" has been done, thus, the lower digits have a slightly greater than 10 percent chance of occurring in the distribution, and the higher digits have a slightly less than 10 percent chance (because of the growth of the population).[13] It is very unlikely that a correction for population growth would, in any case, substantially alter the results. The Myers index is affected by sampling error—i.e., random fluctuating will increase the value of the index—however, values for the index calculated at the national level should not be seriously affected.

Table 3.2 shows the percent distribution of women 20-49 by the last digit of reported age (women 15-44 in Brazil and Guatemala) and the corresponding Myers index.

[13] The blending procedure used by Myers is not appropriate for the limited age range used here.

Table 3.1 Percent distribution of respondents to the individual questionnaire according to information given and imputed for date of birth and current age, Demographic and Health Surveys, 1986-1989

Given:	Yr. & Mo.	Yr. & Age	Age	Year	Yr., Age & Season	Age, Yr. Ignored	None	
Imputed:		Month	Month	Month	Month	Month	Date	
Calculated:..	Age		Year	Age		Year	Age	Total
SUB-SAHARAN AFRICA								
Botswana	84.8	14.6	0.6	--	--	--	a	100.0
Burundi	38.2	61.0	0.7	--	--	--	--	100.0
Ghana	48.7	28.7	3.0	--	--	19.6	--	100.0
Liberia	42.3	21.0	36.7	--	--	--	--	100.0
Mali	9.0	2.5	21.7	b	--	--	--	100.0
Senegal	34.1	22.1	43.8	--	--	--	--	100.0
Togo	26.9	10.1	63.0	--	--	--	--	100.0
Uganda	74.9	25.0	0.1	--	--	--	--	100.0
Zimbabwe	89.9	7.8	2.3	--	--	--	--	100.0
NORTH AFRICA								
Morocco	12.9	76.7	7.4	a	3.0	--	--	100.0
Tunisia	94.2	3.4	2.4	--	--	--	--	100.0
ASIA								
Indonesia	48.5	51.2	0.3	--	--	--	--	100.0
Sri Lanka	89.8	6.1	4.1	--	--	--	--	100.0
Thailand	88.7	10.0	1.4	--	--	--	--	100.0
LATIN AMERICA/CARIBBEAN								
Brazil	99.4	--	0.6	--	--	--	--	100.0
Colombia	98.9	0.8	0.3	--	--	--	--	100.0
Dominican Republic	100.0	--	--	--	--	--	--	100.0
Ecuador	96.9	3.1	a	--	--	--	--	100.0
Guatemala	96.3	3.0	0.7	--	--	--	--	100.0
Mexico	98.0	1.9	0.1	--	--	--	--	100.0
Peru	99.0	0.7	--	0.2	--	--	a	100.0
Trinidad & Tobago	99.8	0.1	0.1	--	--	--	--	100.0

[a]Less than 0.1 percent
[b]In Mali, 66.8 percent had both year and age (age taken as correct in 34.6 percent and year taken as correct in 32.2 percent).

Table 3.2 Percent distribution of women 20-49 by reported terminal digit of age (individual questionnaire) and Myers index, Demographic and Health Surveys, 1986-1989

Country	Terminal Digit 0	1	2	3	4	5	6	7	8	9	Myers Index
SUB-SAHARAN AFRICA											
Botswana	13.5	10.7	11.2	8.6	10.6	10.0	10.0	7.5	10.4	7.6	12.7
Burundi	13.6	8.7	11.4	8.9	9.0	12.3	8.1	10.4	9.6	8.1	15.3
Ghana	17.9	6.8	10.9	7.9	8.0	15.2	9.5	6.4	11.4	6.1	30.6
Liberia	16.5	6.9	10.4	7.8	7.5	13.0	10.4	7.9	11.2	8.3	23.0
Mali	14.0	9.3	10.0	8.2	9.4	10.9	11.6	8.1	10.2	8.4	13.3
Senegal	12.8	10.4	11.7	9.7	9.2	11.7	10.8	8.3	8.3	7.0	14.9
Togo	14.6	9.9	10.8	9.9	8.0	14.2	8.2	8.6	9.7	6.1	19.1
Uganda	15.1	9.6	10.4	9.6	9.9	11.7	10.1	6.0	10.9	6.8	16.3
Zimbabwe	13.5	9.8	12.6	9.1	9.9	9.3	9.8	7.5	9.0	9.5	12.1
NORTH AFRICA											
Morocco[a]	10.9	8.6	10.2	8.9	9.2	10.4	10.1	12.4	9.3	10.1	8.2
Tunisia[a]	9.9	9.0	10.2	10.4	10.6	10.9	10.0	9.4	10.7	9.0	5.6
ASIA											
Indonesia[a]	11.5	8.7	10.7	9.6	8.7	13.9	9.1	10.8	9.1	7.9	13.8
Sri Lanka[a]	8.9	8.8	10.5	10.0	11.4	9.9	10.1	9.6	11.3	9.5	6.7
Thailand[a]	10.2	9.8	9.9	9.5	10.5	10.3	10.6	10.2	9.7	9.4	3.5
LATIN AMERICA/CARIBBEAN											
Brazil[b]	10.4	8.9	9.5	9.6	8.4	11.5	11.4	10.4	9.9	10.1	7.5
Colombia	12.6	10.3	11.2	10.0	10.4	9.5	10.3	8.7	9.0	8.1	9.6
Dominican Republic	14.6	10.5	10.4	11.0	9.9	9.7	9.3	8.6	8.6	7.5	12.8
Ecuador	12.3	10.7	10.8	10.2	10.0	11.3	9.9	8.8	8.7	7.4	10.6
Guatemala[b]	11.1	7.9	9.2	9.0	7.4	12.6	11.4	12.0	10.4	9.0	14.9
Mexico	12.3	10.0	10.0	10.8	10.4	9.8	9.5	8.9	9.4	8.9	7.1
Peru	12.8	10.7	11.5	9.5	9.6	10.1	9.7	9.3	9.4	7.6	10.2
Trinidad & Tobago	10.3	11.7	11.0	9.6	10.5	10.2	9.3	9.5	9.6	8.3	7.5

[a]Ever-married women
[b]Based on women 15-44

The table below ranks the DHS surveys in terms of the extent of digit preference from low to high, according to values of the Myers index.

Digit preference based on the Myers index (M)

Low (M <10)	Moderate (M 10-19)	High (M >20)
Thailand	Peru	Liberia
Tunisia	Ecuador	Ghana
Sri Lanka	Guatemala	
Brazil	Zimbabwe	
Mexico	Botswana	
Trinidad & Tobago	Dominican Republic	
Morocco	Mali	
Colombia	Indonesia	
	Senegal	
	Burundi	
	Uganda	
	Togo	

3.3 FIVE-YEAR AGE GROUP DISTORTIONS

While age heaping is postulated to arise from equal transfers of persons whose true ages fall on either side of the rounded age or "heap," reports of age can also be more grossly inaccurate. For example, an interviewer having no information about a 32 year old woman other than that she has seven children, estimates her age as 45 years old. Or, an interviewer (or the respondent herself) may systematically estimate an age of twenty or above if childbearing has begun. This type of directional transfer of person from true to reported ages is termed gross age misstatement and may introduce serious biases in estimating rates and proportions based on age groups.[14]

[14] If all age reporting were due simply to age heaping, centering age groups on preferred digits would obviate biases in demographic estimation based on age groups.

Table 3.3 Percent distribution of women by five-year age group (individual questionnaire), Demographic and Health Surveys, 1986-1989

Country	Age Group 15-19	20-24	25-29	30-34	35-39	40-44	45-49	Total	Number of Women
SUB-SAHARAN AFRICA									
Botswana	21.5	21.2	19.4	14.9	10.6	6.6	5.7	100.0	4368
Burundi	18.4	19.6	20.2	15.7	12.2	6.8	7.0	100.0	3970
Ghana	18.9	19.3	19.3	14.3	11.8	8.1	8.2	100.0	4488
Liberia	21.7	19.7	20.6	12.6	11.9	6.2	7.3	100.0	5239
Mali	16.3	16.6	19.5	16.1	13.3	9.9	8.3	100.0	3200
Senegal	22.1	20.3	19.0	14.9	10.9	6.8	6.1	100.0	4415
Togo	21.5	19.7	18.5	14.3	11.0	7.8	7.2	100.0	3360
Uganda	24.5	20.8	18.2	13.1	9.7	7.3	6.4	100.0	4730
Zimbabwe	24.3	20.0	16.2	14.0	11.0	7.6	6.9	100.0	4201
NORTH AFRICA									
Morocco[a]	24.6	19.6	17.5	13.1	10.1	7.3	7.8	100.0	5982
Tunisia[a]	20.5	21.2	17.1	15.0	11.0	8.3	6.2	100.0	4184
ASIA									
Indonesia[a]	21.9	19.2	17.1	13.3	10.5	8.7	9.2	100.0	11884
Sri Lanka[a]	20.2	17.8	16.8	14.8	13.1	10.3	6.9	100.0	5865
Thailand[a]	19.4	19.0	16.9	15.2	12.0	9.2	8.3	100.0	6775
LATIN AMERICA/CARIBBEAN									
Brazil	22.1	19.8	17.8	15.9	13.4	11.1	--	100.0	5892
Colombia	22.7	20.3	17.0	13.8	11.1	8.0	7.0	100.0	5331
Dominican Republic	25.6	21.7	16.0	12.7	10.1	7.5	6.4	100.0	7645
Ecuador	22.1	19.5	17.7	14.1	12.0	8.4	6.2	100.0	4713
Guatemala	22.9	19.7	18.7	15.4	13.7	9.6	--	100.0	5160
Mexico	23.8	18.0	17.2	13.9	11.2	8.9	7.0	100.0	9310
Peru	21.9	18.9	16.2	13.6	11.7	9.7	7.9	100.0	4999
Trinidad & Tobago	17.9	19.6	19.6	14.3	11.6	9.7	7.3	100.0	3806

Note: Estimates for samples of ever-married women were calculated using expansion factors based on the proportion of women ever married by age group. Percents may not add to 100 due to rounding.

[a]Ever-married women

Gross misstatement of age can be detected by internal comparison. This is done by distributing respondents by five-year age group, then comparing age ratios between groups. There are some difficulties in using this procedure, however. The dearth of women at the boundaries of the eligible age range (due to the exclusion discussed above), especially at the upper boundary, makes it hard to detect if women have been transferred into the neighboring age group, especially the 40-44 group, since the number of women in the 45-49 group will be too low.

The detection of age-group transference is problematic for ever-married samples because the age distribution for ever-married women under age 24 does not follow an established pattern. To avoid this problem, the age distribution of women of all marital statuses has been estimated based on the total number of respondents. This is done by dividing the number of respondents by the proportion ever-married by age, obtained from the household questionnaire. This is not a true evaluation of the respondents' (mis)reports from the individual survey, because it incorporates elements of misreporting of age and marital status from the household survey. Since most of the estimates calculated for data from the individual questionnaire are based on all women,[15] evaluating these estimates should be done in any case.

Table 3.3 presents the age distributions for data from the individual questionnaire according to five-year age group. The distributions are shown graphically in Figure 3.1.

The data would be expected to show a roughly monotonic decrease in the percentages with increasing age. Six of the 22 surveys present substantial deviations from the expected pattern: Burundi, Ghana, Liberia, Mali, Tunisia, and Trinidad and Tobago. Seven other surveys (Botswana, Senegal, Togo, Zimbabwe, Morocco, Indonesia, and Thailand) appear to have too few women age 40-44 compared with the

[15] The same expansion factors (based on proportion ever married) used in this analysis, are commonly used to estimate other demographic parameters.

Figure 3.1 Age distribution of women based on data from the individual questionnaire, Demographic and Health Surveys, 1986-1989

SUB-SAHARAN AFRICA

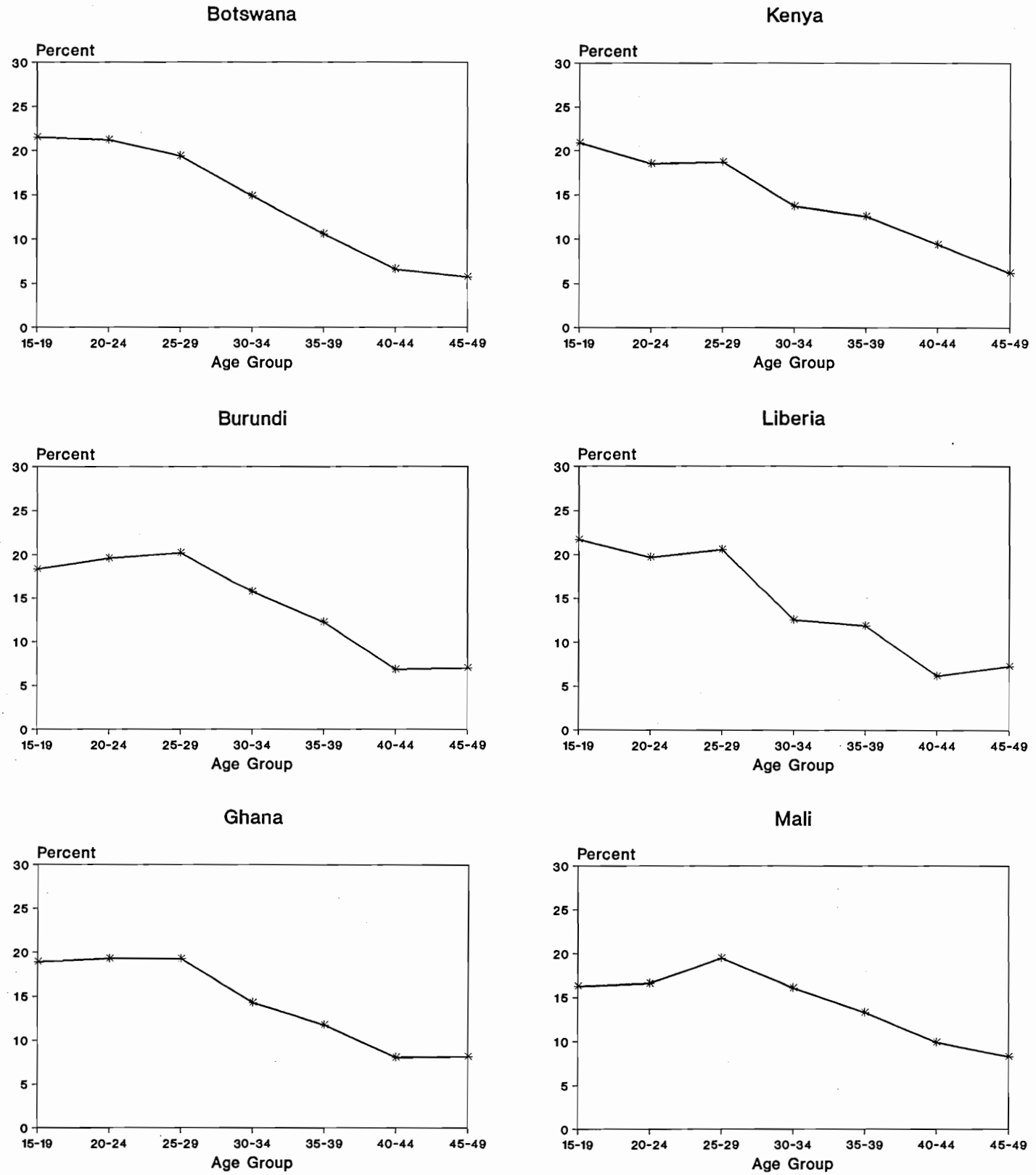
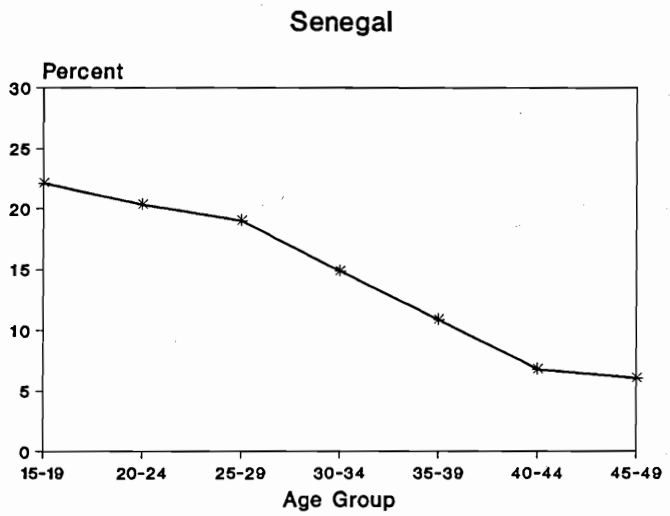
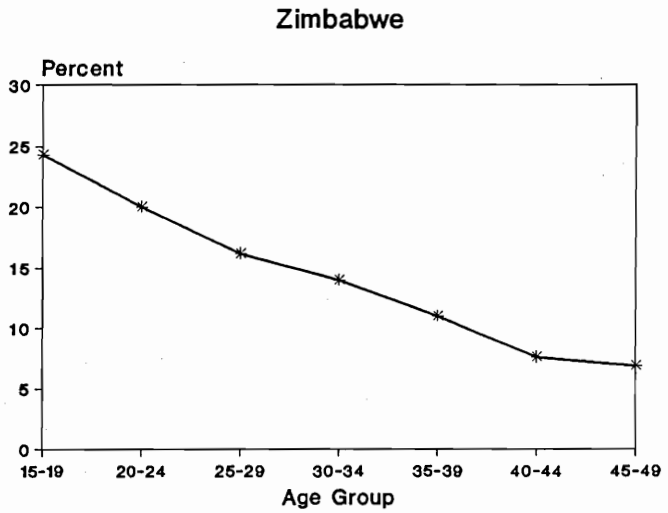
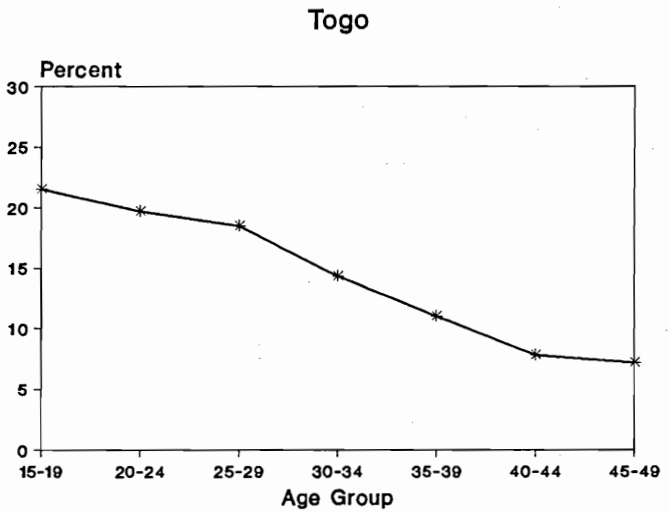


Figure 3.1—*Continued*







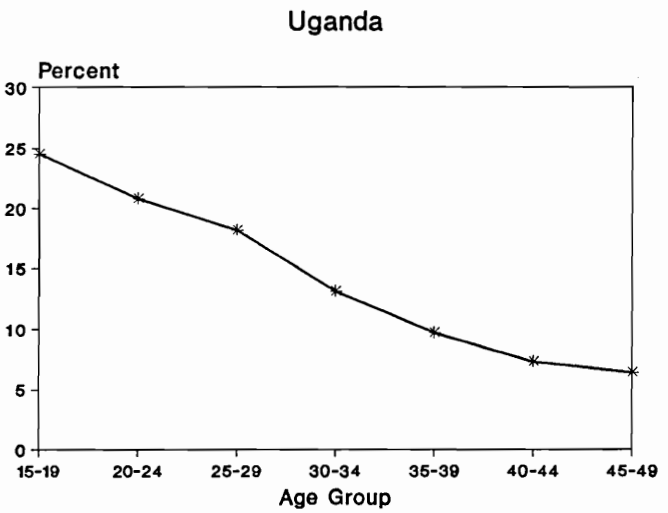
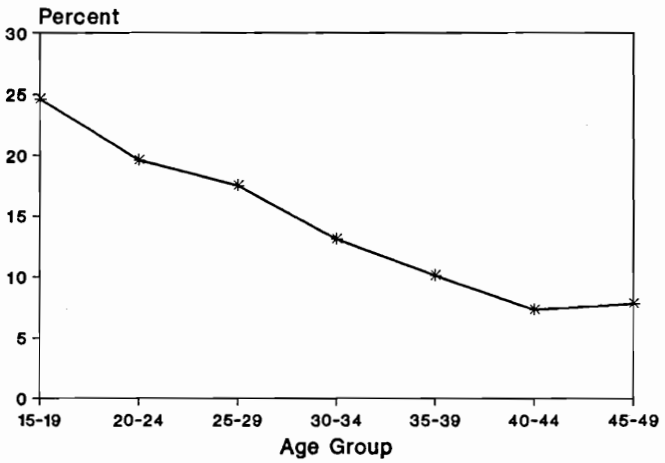


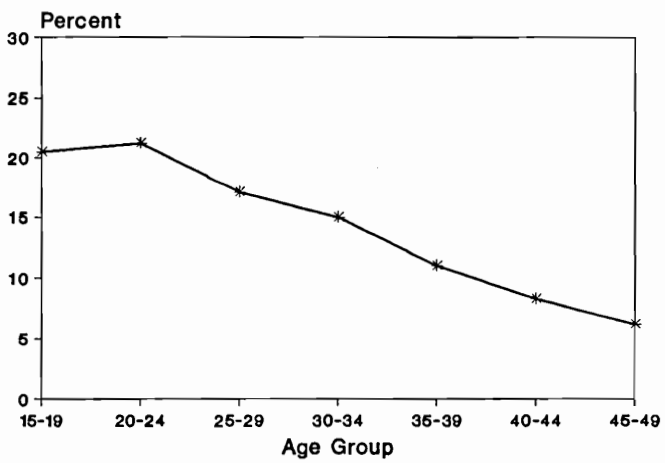
Figure 3.1—*Continued*

NORTH AFRICA

Morocco

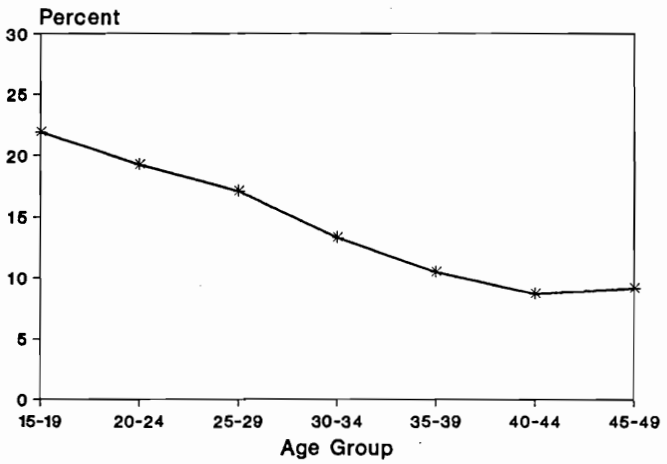


Tunisia

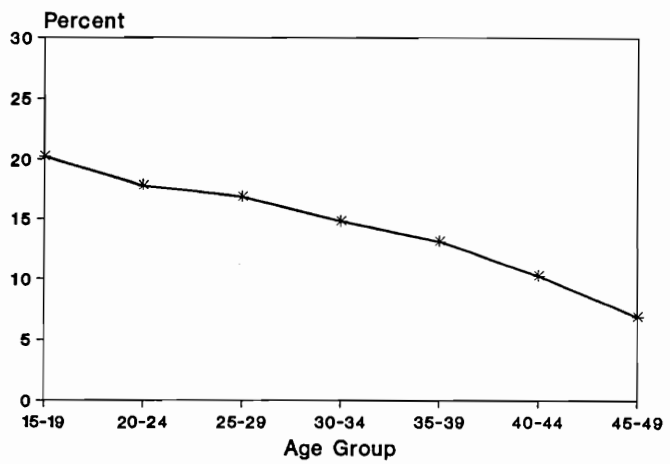


ASIA

Indonesia



Sri Lanka



Thailand

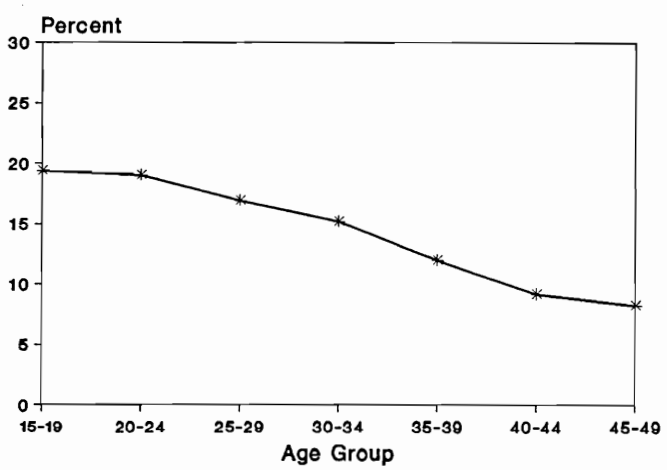


Figure 3.1—*Continued*

LATIN AMERICA/CARIBBEAN

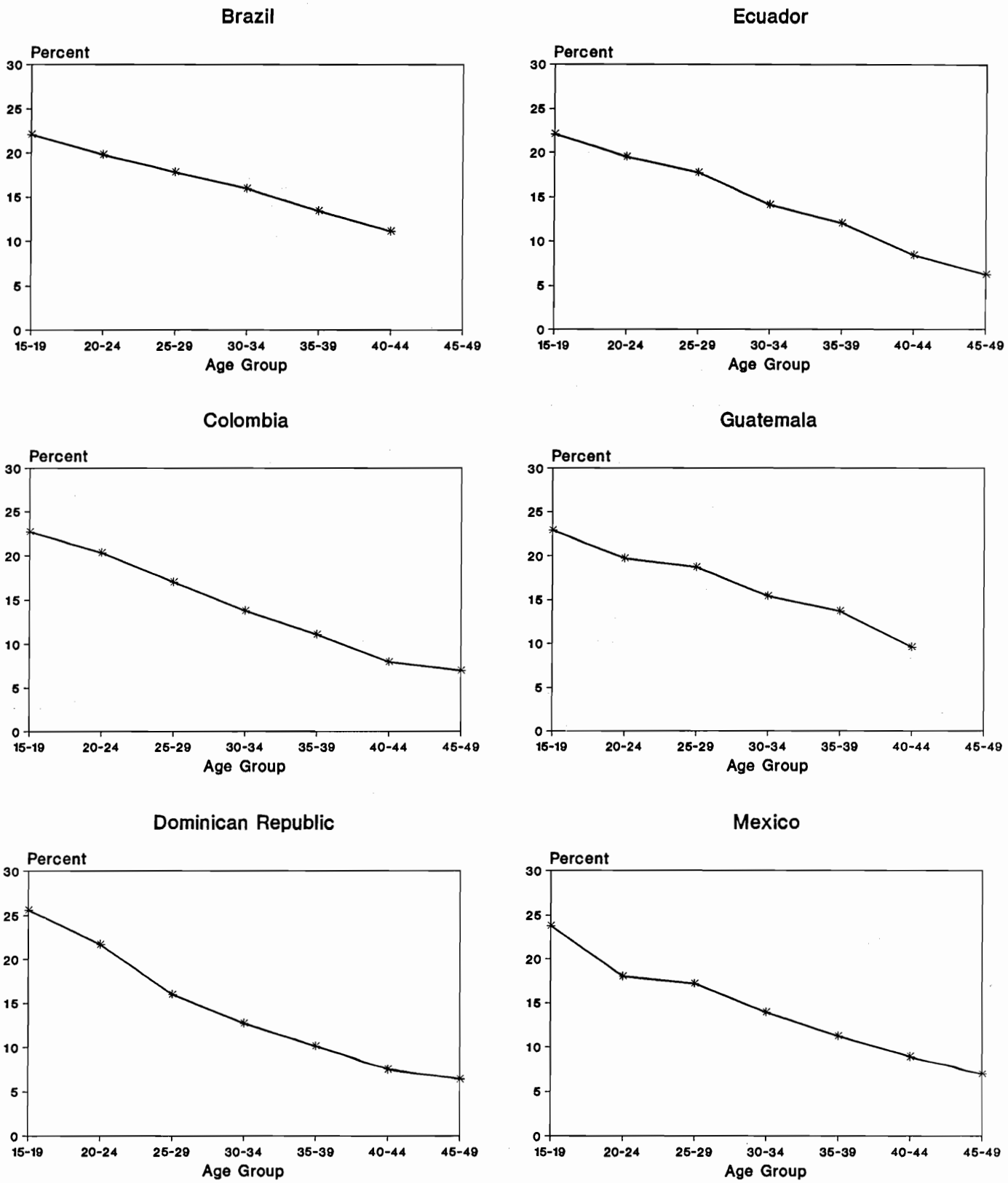
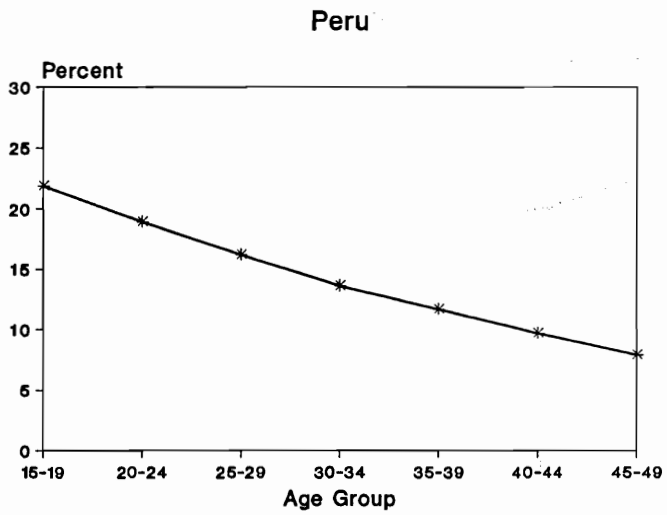
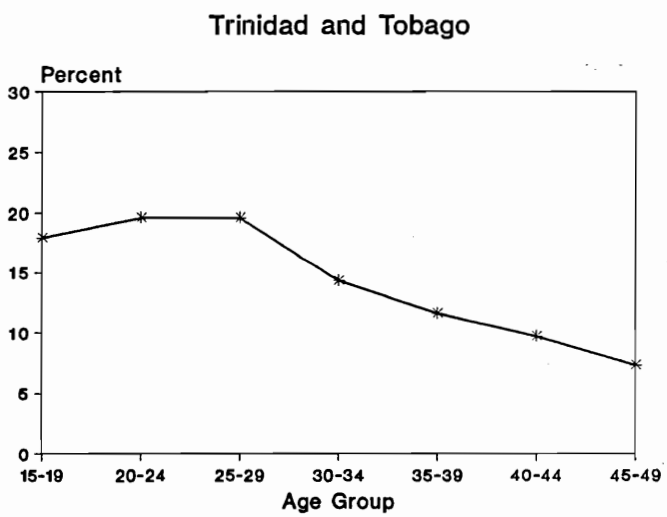


Figure 3.1—*Continued*





number in the 45-49 age group—this, despite in-transference over the 49-year boundary. While some of the deviations from the expected pattern may be due to real variations in the distributions (because of war, famine, and international migration), it is probable that most are due to errors in age reporting.

The table below shows the five-year age groups for which there appears to be a deficit or surplus of women.

Five-year age groups for which there is a deficit or surplus of women

Country	Deficit Groups	Surplus Groups
SUB-SAHARAN		
Botswana	15-19, 40-44	25-29
Burundi	15-19, 40-44	25-29, 30-34
Ghana	15-19, 40-44	25-29
Liberia	15-19, 30-34, 40-44	25-29, 45-49
Mali	15-19, 20-24	25-29, 30-34, 35-39
Senegal	40-44, 45-49	20-24, 25-29, 30-34
Togo	15-19, 40-44	25-29
Uganda	40-44, 45-49	25-29
Zimbabwe	40-44	15-19
NORTH AFRICA		
Morocco	40-44	15-19, 25-29
Tunisia	15-19, 45-49	20-24, 30-34
ASIA		
Indonesia	35-39	25-29, 45-49
Sri Lanka	15-19, 20-24, 45-49	35-39
Thailand	15-19, 40-44	30-34, 45-49
LATIN AMERICA/CARIBBEAN		
Brazil	None	None
Colombia	None	None
Dominican Rep.	40-44, 45-49	20-24
Ecuador	40-44, 45-49	25-29
Guatemala	40-44,	25-29, 35-39
Mexico	20-24	15-19
Peru	None	None
Trinidad & Tob.	15-19	25-29

In order to measure the amount of distortion in the five-year age group distributions, age ratios have been calculated (see Table 3.4). Two indices were created from the age ratios: the sums of the absolute value of the deviations of the ratios from 100 for (1) age groups 20-24 to 40-44 and (2) age groups 25-29 to 40-44. The values of the sums represent the amount of distortion present in the age group distribution.

In the table below the surveys have been classified—using the first index of age group distortion (the sum for age groups 20-24 to 40-44)—into those with values less than twenty (low), 20 to 35 (moderate) and more than 35 (high). Within categories, the surveys are ranked according to increasing value on the index.

Distortion in the five-year age group distribution based on the age-group distortion index (I)

Low (I <20)	Moderate (I 20-35)	High (I >35)
Brazil[a]	Thailand[b]	Tunisia[b]
Peru	Dominican Republic	Morocco[b]
Sri Lanka[b]	Mexico	Mali
Colombia	Indonesia[b]	Ghana
Ecuador	Burundi	
Zimbabwe	Liberia	
	Togo	
	Uganda	
	Senegal	
	Botswana	
	Guatemala[a]	
	Trinidad & Tobago	

[a]Index adjusted for missing 40-44 age ratio
[b]Estimated age distribution using proportions ever married from the household survey

3.4 COMPARISON WITH THE WFS SURVEYS

Table 3.4 compares the DHS and WFS surveys according to values of the age-group distortion index (age range, 25-29 to 40-44). In five of the 13 countries where both DHS and WFS surveys have been conducted, the data indicate substantial improvements in age reporting—the DHS value is ten or more points lower than the WFS value. In two countries, Ecuador and Ghana, the DHS value is ten or more points higher, indicating that age reporting was better in the WFS survey than in the DHS survey.

Table 3.4 Age ratios for five-year age groups, Demographic and Health Surveys, 1986-1989, and comparison of the age-group distortion index for DHS and WFS surveys

	Age Ratios					Age-Group Distortion Index		
						DHS		WFS
Country	20-24	25-29	30-34	35-39	40-44	20-44	25-44	25-44
SUB-SAHARAN AFRICA								
Botswana	103.9	107.1	99.8	98.3	81.3	31.7	27.8	--
Burundi	101.6	114.7	96.5	109.1	70.3	58.5	56.9	--
Ghana	101.0	114.8	92.1	105.4	81.2	47.9	46.9	9.6
Liberia	92.9	128.0	77.2	127.1	65.0	120.1	113.0	--
Mali	92.5	119.2	98.2	102.3	91.8	39.0	31.5	--
Senegal	98.7	108.1	99.5	100.4	79.9	30.3	29.0	36.2
Togo	98.4	108.8	96.8	99.7	85.8	28.1	26.5	--
Uganda	97.7	107.0	94.2	95.0	90.5	29.6	27.3	--
Zimbabwe	98.8	95.0	103.1	102.3	84.4	27.2	26.0	--
NORTH AFRICA								
Morocco[a]	93.1	107.1	95.0	99.0	81.4	38.6	31.7	31.1
Tunisia[a]	112.3	94.8	106.2	95.7	92.4	35.6	23.3	29.4
ASIA								
Indonesia[a]	98.7	105.1	96.4	95.4	88.6	26.1	24.8	23.9
Sri Lanka[a]	95.9	103.3	98.8	104.3	103.1	16.0	11.9	30.8
Thailand[a]	105.0	99.1	104.7	98.4	91.2	20.9	15.9	30.0
LATIN AMERICA/CARIBBEAN								
Brazil	99.2	99.6	101.6	99.4	--	4.3[b]	3.8	--
Colombia	102.2	99.8	98.2	101.9	87.8	18.3	16.1	29.0
Dominican Republic	104.2	93.1	97.2	100.2	90.8	23.2	19.0	59.9
Ecuador	98.1	105.3	94.8	107.0	92.3	27.1	25.2	10.9
Guatemala	94.5	106.8	94.7	109.9	--	34.4[b]	28.9	--
Mexico	88.2	107.7	97.9	98.2	97.5	25.8	14.0	20.4
Peru	98.8	99.8	97.8	100.0	99.2	4.4	3.2	26.1
Trinidad & Tobago	104.3	115.7	91.6	96.6	102.8	34.6	30.3	26.7

Note: Age ratios for samples of ever-married women were calculated using expansion factors based on the proportion of women ever married by age group.

[a]Ever-married women
[b]Index adjusted for lack of 40-44 age ratio by adding the average observed deviation in age ratios.

4 Summary and Conclusions

From the initial survey and sample design stage, through interviewer training, data collection, and data processing, every effort was made in the DHS surveys to produce the most complete and accurate data possible. Despite such efforts, this analysis indicates that there are weaknesses in the data used to determine eligibility for the individual questionnaire and in the age data collected in the individual questionnaire. The following is a summary of the findings of this analysis, with recommendations for improving data quality in future surveys.

4.1 ELIGIBILITY

In 17 of the 22 surveys examined here, there appears to be substantial exclusion of women at the boundary age groups 15-19 and 45-49 (40-44 for Brazil and Guatemala). The data from six surveys show marked distortion at the lower boundary, indicating out-transference of eligible women; the data from 16 surveys indicate pronounced out-transference (or exclusion) at the upper boundary. When compared with WFS surveys for the same countries, DHS surveys generally show a greater amount of distortion at the eligible age boundaries, especially the upper boundary.

Exclusion from the individual questionnaire due to "sleeping away" the night before is pronounced in six sub-Saharan surveys in which more than five percent of the otherwise eligible women have been excluded. Because the percentage of residents sleeping away is much lower for the age groups just outside the eligible range, it is concluded that many of the women have been misclassified by interviewers so as to lighten their workloads. The conclusion that deliberate exclusion at the age eligibility boundaries has occurred—whether as a result of age misstatement or misclassification of "sleeping-away" status—is reinforced by comparisons of DHS data with data from other sources. These comparisons point to distortions both at the upper and lower boundaries (particularly the upper boundary) that appear to be DHS-specific and cannot be explained by reporting phenomena in the respective populations.

In order to estimate whether such exclusions may have biased the demographic rates, a series of simulations were conducted under several sets of assumptions. The results indicated that only under the most extreme assumptions concerning the demographic behavior of "excluded women" would estimates for the total fertility rate and under-five mortality be biased. The most problematic cases involved surveys in sub-Saharan Africa—especially Botswana, Ghana, Liberia, and Mali—where underestimates for under-five mortality and overestimates for the total fertility rate could be from 3 to 5 percent. The results of the simulation of bias in estimating contraceptive prevalence indicate that exclusion had very little impact—the worst case was the Botswana survey, in which contraceptive prevalence was overestimated by only 2 percent.

4.2 RESPONSE RATES

Household response rates were over 95 percent for all surveys except Brazil (91.9) and Liberia (89.6). For the individual questionnaire, response rates were also very high, with the exception of Brazil (87.5 percent). In view of these uniformly high response rates, unless there are huge differences in the demographic behavior between responding and nonresponding households (and women), there is little possibility that significant bias has occurred in the demographic estimates (except, perhaps, in the case of Brazil).

In summary, one survey (Liberia) rated poorly on all four indices of household data quality (the Myers index, the UN index, the T index combining upper and lower boundary distortions, and the index of exclusion due to "sleeping away" status); four surveys (Botswana, Ghana, Mali and Uganda) rated poorly on three indices; and Burundi, Ecuador, Sri Lanka, and Togo rated poorly on two indices.

4.3 AGE OF RESPONDENT

The quality of age data from the individual questionnaire varies by region. The best age reporting is in Latin America and the Caribbean, where more than 96 percent of women were able to report their year and month of birth. In nine surveys from other regions (seven from sub-Saharan Africa), less than 80 percent of the women interviewed reported a complete birth date—Burundi, Ghana, Indonesia, Liberia, Mali, Morocco, Senegal, Togo, and Uganda.

An analysis of digit preference indicated that there was substantial age heaping in four surveys (Ghana, Liberia, Togo, and Uganda), especially on ages ending in zero and five. Distortions in the five-year age distribution—more important than age heaping per se in calculating many of the demographic rates—were high in six surveys (Burundi, Ghana, Liberia, Mali, Morocco, and Tunisia). Analysts drawing conclusions from these datasets are advised to consider the possibility of bias due to both age heaping and gross age-group transference.

4.4 RECOMMENDATIONS

Because of the comparative nature of this report—which is intended as a reference for users of DHS data—the analysis of the quality of data on eligibility and age is not exhaustive and no attempt has been made to provide explanations for every atypical pattern observed. Ongoing analysis, by DHS staff and researchers worldwide, will likely yield further insights into questions of data quality and contribute to improved methods of data collection.

Since intentional exclusion of women from eligibility was found to be a particular problem in the DHS surveys, the following recommendations have been made in an effort to minimize exclusion in future surveys:

- The question on presence during the previous night should remain in the household questionnaire, but interviewers should interview both resident women and visitors, regardless of whether they spent the previous night in the household; the selection of a de facto or de jure sample should then be made during data processing.

- Supervisors should reinterview a sub-sample of those households where women are reported in age groups just outside the eligible age range, to be sure that the ages of these women were correctly recorded.

- Set the upper limit of age eligibility at 54 years when it is considered especially important to get good information for women 45-49.

- Alternatively, field procedures should be modified so that determination of eligibility is made independently of the individual interview.

Notwithstanding these recommendations, it seems unlikely that exclusion can be completely avoided. In order to be able to establish a profile of excluded women (eligible women not interviewed) in future surveys, basic socioeconomic information on the household and household members should be collected in the household questionnaire.

Most demographic parameters are relatively insensitive to age heaping, which is fortunate since there is probably little that can be done to improve the accuracy of age reporting in countries where there is little documentation of individual age and people attach little significance to knowledge of exact age. However, gross age misstatement can seriously affect estimates derived from survey data. It is in this area that realistic efforts can be made to improve data quality. One possibility is suggested:

- Since an individual's age is first established during the household interview, a more rigorous method of age determination in the household questionnaire could yield better age data. This would involve asking about birth dates and seeking documentation of age.

References

Ewbank, D.C. 1981. *Age Misreporting and Age-Selective Underenumeration: Sources, Patterns and Consequences for Demographic Analysis.* Committee on Population and Demography, Report No. 4. Washington, D.C.: National Academy Press.

Institute for Resource Development. 1987. *Sampling Manual.* Basic Documentation, No. 8. Columbia, Maryland.

Rutstein, Shea Oscar. 1985. "Assessment of the Quality of Age Reports for Eligibility and Analysis." In *Assessment of the Quality of Data in 41 WFS Surveys: A Comparative Approach* by Noreen Goldman, Shea Oscar Rutstein, and Susheela Singh, WFS Comparative Studies, No. 44. Voorburg, Netherlands: International Statistical Institute.

Shryock, Henry S., Jacob S. Siegel, and Associates. 1971. *The Methods and Materials of Demography.* Vol. 1. Washington, D.C.: U.S. Dept. of Commerce, Bureau of the Census.

United Nations. 1967. *Manual II: Methods of Appraisal of Quality of Basic Data for Population Estimates.* Population Studies, No. 23. New York: Dept. of Economic and Social Affairs.

Assessment of the Quality of Data on Age at First Sexual Intercourse, Age at First Marriage, and Age at First Birth in the Demographic and Health Surveys

ANN K. BLANC
NAOMI RUTENBERG

Acknowledgments

The authors wish to express their thanks to Guillermo Rojas for programming assistance. Jeremiah Sullivan and Mary Beth Weinberger provided valuable comments on the manuscript.

1 Introduction

This report examines the quality of the data collected in the Demographic and Health Surveys (DHS) program on the timing of the beginning of the reproductive period in women's lives. Three aspects of the early stages of the reproductive period are examined—age at first sexual intercourse, age at first union, and age at first birth. Accurate information on these three variables is important because they signal the beginning of exposure to the risk of pregnancy and the onset of the childbearing years. At the individual level, a woman's age at the time she begins bearing children is related to the ultimate size of her completed family, as well as to the proportion of her life she devotes to childbearing and childrearing (United Nations, 1988). In addition, early childbearing is associated with an array of health, social, and economic consequences for women and their children (Liskin et al., 1985). The three variables are also of critical interest in the aggregate because they determine the total number of years of exposure to the risk of pregnancy and identify when the period of exposure began.

The conventional marker for the beginning of exposure to the risk of pregnancy is the date of first union. Generally, this date is the beginning of a socially recognized union. In some societies, however, a substantial amount of sexual activity and even childbearing may precede the recognized union. Thus, even when the date of a first union is accurately reported, some exposure may be missed. In an attempt to improve the estimates of exposure to pregnancy, especially among teenagers, the DHS questionnaire includes a question on the age at which sexual intercourse first occurred.

The misreporting of date of first union or age at first sexual intercourse has an impact on the analysis of changes in the age at which women are first exposed to pregnancy and on fertility rates based on total exposure time. Misreporting of first (or later) birth dates can lead to inflated estimates of recent fertility decline (or apparent decline) when, in fact, fertility has remained constant. A common type of error in data from developing countries is forward displacement of early events by older respondents; older women tend to report early events as occurring closer to the survey date (i.e., at an older age) than they actually occurred. The combination of forward displacement of early births with accurate reporting of recent births results in a concentration of births in an intermediate period (e.g., 5-15 years ago). The presence of such errors can lead to incorrect fertility estimates, which give the appearance of increasing fertility in earlier periods, e.g. 20-30 years ago, and decreasing fertility in recent periods (Goldman, 1985).

This report presents the results of three tests used to assess the extent of errors in the reporting of date of first union and first birth: 1) the completeness of reporting of dates and selected indicators of heaping, 2) internal checks of consistency, and 3) comparisons of DHS data with data from other surveys and censuses, where possible. The evaluation of reported age at first sexual intercourse is based entirely on checks of internal consistency, as this information has rarely been collected in other surveys.

Standard recode data files were used for all tabulations included in this report. The editing and imputation procedures applied in the creation of the DHS datasets are described below. While the procedures described herein are standardized for the DHS program, it should be noted that individual countries have introduced some modifications (usually minor) in the editing guidelines.

DHS data are edited (cleaned) in four stages. First, the field editors or supervisors review completed questionnaires in the field. Any inconsistencies discovered—for example, disagreement between the date of birth of a child and the age of the child—are resolved by the interviewer and editor. If the inconsistency cannot be rectified based on the interviewer's recall or other information in the questionnaire, the interviewer is sent back to the respondent to obtain the correct information. Second, the data entry programs used in DHS surveys are designed to ensure that the data entered are free of basic errors such as values that are out of range. Additionally, any remaining discrepancies between dates and ages are corrected at this time.

Next, a secondary editing program checks for further logical inconsistencies in the data—for example, if the first birth occurred when the respondent was less than 12 years old.[1] Errors of this type are either corrected after a careful review of the questionnaire or are given a special code (97), which indicates that they are inconsistent with some other piece of information in the questionnaire. The dataset is preserved at this stage and is referred to as the "raw" data.

The final step in editing the data is machine imputation. The method used in DHS for imputation of dates relies on the construction of logical ranges for each date; these are

[1] A comprehensive description of secondary edits can be found in the DHS Data Processing Manual (Institute for Resource Development, 1989).

refined in a series of steps resulting in successively narrower or constrained ranges. All missing dates—due to the interviewer not recording the information, the respondent not knowing the information, or the date being changed to "inconsistent" during secondary editing—are imputed, including the date of first union and the date of first birth. Since information on the timing of first sexual intercourse is collected in the form of age (in years) rather than a specific date, it is not imputed.

The outcome of the imputation procedure is a complete date of first union for all ever-married women and a complete date of first birth (and subsequent births) for women who gave birth. The age at first union or the current age of the child is then calculated based on the imputed date. Each record is flagged with a variable that describes which data have been reported, and by implication, which data it was necessary to impute. The dataset with the imputed dates (which is also preserved) is the basis of the standard recode files. At the time of this analysis, standard recode files were available for 22 countries in Africa, Asia, Latin America, and the Caribbean. The sample size and year of fieldwork for these 22 surveys are shown in Table 1.1.

Table 1.1 Background information and type of data collected on union status, date of first union, and age at first union, Demographic and Health Surveys, 1986-1989

				Type of Data		
Country	Date of Fieldwork	Sample Size	Sample Type	Current Union Status	Date of First Union	Question on Age at First Sexual Intercourse
SUB-SAHARAN AFRICA						
Botswana	1988	4368	All	A	A	Yes
Burundi	1987	3970	All	C	A	Yes
Ghana	1988	4488	All	A	A	Yes
Liberia	1986	5239	All	A	A	Yes
Mali	1987	3200	All	A	A	Yes
Senegal	1986	4415	All	B, C	A	No
Togo	1988	3360	All	A	E	Yes
Uganda	1988/89	4730	All	A	A	Yes
Zimbabwe	1988/89	4201	All	F	A	Yes
NORTH AFRICA						
Morocco	1987	5982	EM	A	A	No
Tunisia	1988	4184	EM	A	A	No
ASIA						
Indonesia	1987	11884	EM	A	D	Yes
Sri Lanka	1987	5865	EM	A	A	No
Thailand	1987	6775	EM	A	D	Yes
LATIN AMERICA/CARIBBEAN						
Brazil	1986	5982	All	B	A	Yes
Colombia	1986	5329	All	A	A	Yes
Dominican Republic	1986	7649	All	A	D	Yes
Ecuador	1987	4713	All	A	D	Yes
Guatemala	1987	5160	All	A	A	Yes
Mexico	1987	9308	All	B	G	Yes
Peru	1986	4999	All	A	D	Yes
Trinidad & Tobabo	1987	3806	All	H	A	Yes

EM = Ever-married women
All = All women

A = Standard core questions or equivalent information
B = Probe on ever lived together added for never-married women
C = Probe on currently living together added for women formerly in union
D = Date of first union AND age at first union asked of all ever-married women
E = Marriage history added
F = "Living together" category excluded
G = Question on age at first marriage excluded; probe on living together before marriage added.
H = Visiting unions added

2 Date of First Union

In the DHS surveys, marriage is of interest primarily as an indicator of the beginning of exposure to pregnancy; it is defined as including all stable sexual relationships regardless of the legal status of the union. Thus, women in both formal and informal unions are considered "in union" or married. Although most DHS surveys include all women of reproductive age irrespective of marital status, five of the countries in this report—Indonesia, Morocco, Sri Lanka, Thailand, and Tunisia—included only ever-married women. In spite of the broad definition of marriage used in DHS surveys, cultural differences in the way unions are formed, their social significance, and their relationship to childbearing may affect the way union status and the timing of first unions are reported.

In the DHS model questionnaire, information on marriage is collected in a series of questions. Women are first asked whether they have ever been married or have lived with a man. Those who report having had a husband or partner are then asked about their current union status and whether they have been in one or more than one union. Finally, respondents are asked to provide the month and year they started living with their first husband or partner. If the woman cannot provide a year of first union, she is then asked to give her age at the time of the first union.

Local variations in union types and formation, as well as interest in maintaining consistency with previous surveys, prompted many countries to make modifications in the standard set of questions. These modifications are summarized in Table 1.1. Three countries (Brazil, Mexico, and Senegal) added a probe for never-married women, on whether the respondent ever lived with a partner; two countries (Burundi and Senegal) added a probe for women who declared their current union status was widowed, divorced, or separated, on whether the respondent was currently living with a partner. One country, Zimbabwe, dropped the category "living together," while Trinidad and Tobago added the category "visiting union."[2]

The principal modification to the recommended data collection procedure for date of first union was to ask all ever-married women for both the date of the beginning of their first union and their age at that time. This alteration was incorporated into the questionnaires used in the Dominican Republic, Ecuador, Indonesia, Peru, and Thailand.[3] During editing, survey staff found that a substantial number of respondents reported a date of first union that was inconsistent with the reported age at first union. To reconcile the inconsistencies, subjective and largely unrecorded decisions were made in each case, based on the survey staff's opinion about which was the more reliable piece of information. These inconsistencies suggest that women may find it difficult to provide precise information about their first union and that both date of first union and age at first union are often the respondent's best guess.

The Mexican survey asked only for the date of first union, omitting the question on age at first marriage. This implies that interviewers had to probe more intensively in Mexico as they were required to collect a complete date for the first union from every respondent.

2.1 COMPLETENESS OF INFORMATION

Table 2.1 shows the proportion of respondents by age group who reported a complete date of first union (i.e., year and month) and the extent of imputation necessary in the remaining cases. The columns in Table 2.1 are ordered from left to right according to the degree of imputation. Column one shows the proportion of cases for which a complete date was reported. Column two is the percentage of cases for which a year was reported but a month was imputed. Column three indicates the percentage of cases for which only age at marriage was reported and the year and month were imputed based on that age. Column four shows the percentage of cases for which no information was provided and both the year and month were imputed based on the dates of surrounding events.

The most complete date reporting occurs in Latin America and the Caribbean, followed by Asia, and North, East, and Southern Africa; the poorest reporting comes from West Africa. In seven countries, predominantly in Latin America, more than 80 percent of ever-married women reported

[2] Women in union in Trinidad and Tobago include those in visiting unions as well as formal and common-law marriages.

[3] For these surveys, the collection of date of first union and age at first union was most likely inadvertent. An early version of the model questionnaire did follow this flow, but was subsequently changed. Nonetheless, two early surveys–Peru and Thailand–adapted their questionnaires from this version of the model questionnaire. Subsequently, the questionnaires used in the Dominican Republic and Ecuador were modified from the Peru questionnaire, and the Indonesia questionnaire was adapted from the Thailand questionnaire, perpetuating the inclusion of both questions.

Table 2.1 Percent distribution of ever-married women by completeness of information on date of first union, by age group, Demographic and Health Surveys, 1986-1989

Country	Year and Month Reported, No Imputation	Year Reported, Month Imputed	Age Reported, Year and Month Imputed	No Information Year and Month Imputed	Total
SUB-SAHARAN AFRICA					
Botswana					
20-24	76.2	22.6	0.5	0.7	100.0
30-34	64.6	30.8	1.8	2.7	100.0
45-49	60.6	33.8	3.2	2.5	100.0
Total	65.7	30.8	1.5	2.0	100.0
Burundi					
20-24	78.6	20.3	1.1	0.0	100.0
30-34	53.2	43.7	3.1	0.0	100.0
45-49	38.8	55.3	5.4	0.4	100.0
Total	57.7	39.5	2.7	0.1	100.0
Ghana					
20-24	35.8	53.6	9.7	0.9	100.0
30-34	28.1	56.3	13.5	2.0	100.0
45-49	19.1	55.2	24.0	1.6	100.0
Total	29.3	55.3	13.8	1.5	100.0
Liberia					
20-24	37.2	45.6	15.1	2.2	100.0
30-34	28.2	37.3	33.2	1.3	100.0
45-49	14.9	22.6	57.8	4.8	100.0
Total	29.0	38.4	30.7	1.9	100.0
Mali[a]					
20-24	9.3	23.4	44.2	22.4	a
30-34	6.4	12.6	52.3	28.7	a
45-49	0.2	5.4	70.6	23.7	a
Total	5.9	17.7	50.7	25.0	a
Senegal					
20-24	19.0	9.5	69.4	2.0	100.0
30-34	14.3	8.9	75.6	1.2	100.0
45-49	8.5	10.7	80.4	0.4	100.0
Total	16.6	9.2	72.2	2.0	100.0
Togo					
20-24	31.1	23.6	44.9	0.4	100.0
30-34	15.6	19.4	64.3	0.6	100.0
45-49	4.5	12.3	81.1	2.1	100.0
Total	19.6	20.2	59.4	0.9	100.0
Uganda					
20-24	91.4	7.6	0.9	0.1	100.0
30-34	86.0	10.6	2.6	0.6	100.0
45-49	69.5	23.4	6.6	0.5	100.0
Total	86.0	11.2	2.3	0.5	100.0
Zimbabwe					
20-24	88.0	10.3	1.5	0.2	100.0
30-34	76.5	18.5	4.7	0.4	100.0
45-49	58.4	30.4	11.2	0.0	100.0
Total	76.9	17.8	5.0	0.3	100.0

Table 2.1-*Continued*

Country	Year and Month Reported, No Imputation	Year Reported, Month Imputed	Age Reported, Year and Month Imputed	No Information Year and Month Imputed	Total
NORTH AFRICA					
Morocco[b]					
20-24	28.1 (60.0)	3.4	8.3	0.2	100.0
30-34	24.2 (62.2)	5.8	7.6	0.2	100.0
45-49	7.1 (63.6)	15.4	13.3	0.5	100.0
Total	23.4 (60.0)	7.2	9.2	0.2	100.0
Tunisia					
20-24	76.0	13.3	10.7	0.0	100.0
30-34	62.3	19.6	18.0	0.1	100.0
45-49	43.5	18.4	37.9	0.2	100.0
Total	61.8	18.0	20.0	0.1	100.0
ASIA					
Indonesia					
20-24	84.7	15.3	0.1	0.0	100.0
30-34	71.6	28.0	0.3	0.0	100.0
45-49	45.0	54.0	0.4	0.6	100.0
Total	68.2	31.3	0.3	0.2	100.0
Sri Lanka					
20-24	83.9	10.6	5.4	0.1	100.0
30-34	81.4	13.0	5.6	0.0	100.0
45-49	64.8	17.6	17.7	0.0	100.0
Total	78.4	13.3	8.3	0.0	100.0
Thailand					
20-24	84.1	13.8	0.0	2.1	100.0
30-34	75.5	21.9	0.0	2.6	100.0
45-49	66.5	30.1	0.0	3.4	100.0
Total	75.8	20.9	0.0	3.2	100.0
LATIN AMERICA/CARIBBEAN					
Brazil					
20-24	92.9	4.8	2.3	0.0	100.0
30-34	90.5	5.1	4.4	0.0	100.0
40-44	90.4	4.4	5.2	0.0	100.0
Total	91.4	4.4	4.1	0.0	100.0
Colombia					
20-24	91.5	5.6	2.9	0.0	100.0
30-34	87.2	8.9	3.9	0.0	100.0
45-49	82.6	11.7	5.7	0.0	100.0
Total	87.9	8.1	4.0	0.0	100.0
Dominican Republic					
20-24	88.2	11.2	0.7	0.0	100.0
30-34	80.9	18.0	1.0	0.0	100.0
45-49	67.3	29.3	3.4	0.0	100.0
Total	81.2	17.5	1.3	0.0	100.0
Ecuador					
20-24	88.1	9.9	1.8	0.2	100.0
30-34	83.3	13.4	2.8	0.5	100.0
45-49	68.2	25.4	5.3	1.0	100.0
Total	82.3	14.4	2.7	0.6	100.0

Table 2.1—*Continued*

Country	Year and Month Reported, No Imputation	Year Reported, Month Imputed	Age Reported, Year and Month Imputed	No Information Year and Month Imputed	Total
Guatemala					
20-24	85.4	9.0	0.0	5.6	100.0
30-34	75.6	12.8	0.0	11.6	100.0
40-44	65.2	21.3	0.0	13.5	100.0
Total	77.8	13.3	0.0	8.9	100.0
Mexico					
20-24	96.0	2.9	0.0	1.0	100.0
30-34	93.9	4.7	0.0	1.3	100.0
45-49	87.1	10.0	0.0	2.3	100.0
Total	93.5	5.0	0.0	1.4	100.0
Peru					
20-24	95.2	3.5	1.3	0.0	100.0
30-34	90.1	7.8	2.1	0.0	100.0
45-49	75.5	16.0	8.4	0.0	100.0
Total	87.9	8.9	3.2	0.0	100.0
Trinidad & Tobago					
20-24	77.6	18.7	2.8	1.0	100.0
30-34	73.6	22.2	3.5	0.8	100.0
45-49	70.9	23.6	4.4	1.1	100.0
Total	75.4	20.8	2.8	1.0	100.0

Note: The total refers to all women 15-49 (15-44 for Brazil and Guatemala).

[a] Percentages do not sum to 100.0 for date of first union because women whose marriages had not been consummated are excluded.

[b] Numbers in parentheses are the percentage of women who reported a season and year.

the year and month of their first union. Brazil and Mexico have the highest proportion of complete dates, more than 90 percent. Overall, more than 75 percent of women in the Latin American and Caribbean surveys gave the year and month of their first union. Completeness of reporting drops off somewhat in the Asian countries, where between 68 and 78 percent of women provided the date of first union.

Countries in East Africa (Burundi and Uganda), Southern Africa (Botswana and Zimbabwe), and North Africa (Tunisia and Morocco) stand out for having more complete information than countries in West Africa. In these, between 58 and 86 percent of women provided complete information on date of first union.[4] In West Africa, reporting of a complete date of first union was very low in Ghana, Liberia, Mali, Senegal, and Togo, with the poorest reporting in Mali, where only 6 percent reported the year and month of their first union. In many African societies, marriage involves a number of stages and it may be difficult for the respondent to recall when she first began to live with her partner or even when the marriage process began.

In the majority of cases in which the complete date of first union was not specified, either the year of first union or the women's age at first union was available. This information is used in the imputation procedure to constrain the range of months for which a consistent date can be imputed. Thus, in most cases imputed dates are likely to be fairly close to the true dates. When it was necessary to impute part or all of the information for the date of first union, it was most common to impute only the month. A month of first union alone was imputed when a consistent year of the event was reported and, thus, the possible range for the imputation is 12 months. The greatest amount of month imputation occurred in Ghana where month of first union was imputed in more than half of the cases. Other countries in which more than one-quarter of the cases have an imputed month are Botswana, Burundi, Indonesia, and Liberia.

[4] In Morocco, 23 percent reported a year and month and 60 percent reported a season for the date of marriage. For women who reported a season, an exact month of marriage was randomly imputed with a 3-month range for that season.

In a few countries, the majority of women reported only their age at marriage and both the year and month were imputed. If the year but not the month of birth of the respondent is given and age at first union is given, the range for the date of first union is up to 24 months. If the year and month of birth of the respondent are specified and age at first union is reported, the range for imputation is 13 months. In six countries more than 10 percent of women reported only their age at first union: Ghana, Liberia, Mali, Senegal (72 percent), Togo, and Tunisia.

Relatively few women were lacking all three pieces of information—year and month of first union and age at first union—requiring that the date be imputed in a range between some minimum age (usually 12) and the constraint that the first union should begin at least seven months before the birth of the first child. However, 25 percent of the cases in Mali were imputed in this manner and 9 percent in Guatemala. In all other countries, the proportion of women who did not report any information on first union was 3 percent or less.

In general, older women were less able than younger women to report a year and month of first union. For example, in Burundi, 79 percent of women 20-24 reported the year and month of their first union, but only 39 percent of those 45-49 reported a complete date. In only three countries, Brazil, Colombia, and Mexico, did more than 80 percent of the women in the oldest cohort report the year and month of their first union. The large rise in incomplete dates among older women, and the concomitant increase in imputed information, has implications for evaluating trends in age at first marriage.

2.2 HEAPING

Different types of date reporting may lead to different patterns of heaping in the data. When dates are reported in the form of a calendar year, heaping may occur on calendar years ending in zero or five, or in years with notable events. If the respondent had difficulty recalling the date, it is likely that interviewers employed a probe about how long ago the first union began. This line of questioning may have resulted in dates that are heaped on rounded durations of years of marriage. When ages are reported, misreporting may take the form of heaping on rounded ages. This form of heaping is difficult to detect, however, because first marriages are usually highly concentrated in a narrow age range. The extent of imputation will also affect the amount of heaping: since imputation is done randomly within a constrained range (for cases with incomplete data), the effect is to reduce heaping.

The indices of heaping on year of first union or duration since first union ending in zero or five are shown in Table 2.2. A rough indicator of significant heaping is an index over 1.05. Six of the eight Latin American countries (Ecuador and Mexico being the exceptions) and one Asian country (Indonesia) have indices over 1.05 for heaping on years of first union ending in zero or five. It has been

Table 2.2 Indices of heaping on year of first union and duration since first union ending in 0 or 5, Demographic and Health Surveys, 1986-1989

Country	Year of First Union	Duration Since First Union
SUB-SAHARAN AFRICA		
Botswana	1.03	0.94
Burundi	1.09	0.94
Ghana	1.06	0.99
Liberia	1.25	1.20
Mali	0.95	1.09
Senegal	1.05	1.02
Togo	0.96	1.09
Uganda	1.07	0.96
Zimbabwe	1.05	1.01
NORTH AFRICA		
Morocco	1.02	1.01
Tunisia	1.06	1.02
ASIA		
Indonesia	1.11	1.00
Sri Lanka	1.03	0.97
Thailand	1.01	0.99
LATIN AMERICA/CARIBBEAN		
Brazil	1.08	1.05
Colombia	1.08	0.91
Dominican Republic	1.11	1.02
Ecuador	1.01	1.08
Guatemala	1.13	1.06
Mexico	1.03	1.10
Peru	1.13	1.01
Trinidad & Tobago	1.06	0.93

For year of first union:

$$\text{Index} = \sum_{i=1}^{5} \frac{x_i}{(x_{i-2}) + (x_{i-1}) + x_i + (x_{i+1}) + (x_{i+2})}$$

where $i_1 = 1960$, $i_2 = 1965$, $i_3 = 1970$, $i_4 = 1975$, $i_5 = 1980$,
x = number of women married in year i

Index for duration since first union is identical except

$i_1 = 5$, $i_2 = 10$, $i_3 = 15$, $i_4 = 20$, $i_5 = 25$,
x = number of women at duration i since first union

suggested that the heaping of dates of first union in Latin America is due largely to the common occurrence of informal first unions. The exact starting date may be forgotten because of the relatively temporary nature of the union and because of the absence of any social or religious ceremony to mark its beginning (Singh, 1985). Significant heaping on durations of marriage ending in zero or five is found in Ecuador and Mexico, the two countries that do not suffer from heaping on the calendar year of first union. In Mexico, this type of heaping may be a consequence of interviewers probing for the year of union by asking the respondent about the number of years ago the union began.

There is surprisingly little heaping on calendar years or durations ending in zero or five in African countries. The exception is Liberia, which has substantial heaping on both indices, and Mali and Togo, which suffer to a lesser degree from heaping on durations. The lack of heaping, or the small amount of heaping on calendar years in countries such as Mali, Senegal, and Togo, may be due to the large proportion of cases in which the year was imputed on the basis of age at first union, with the result that a considerable degree of randomness was introduced into the distribution of dates. Dates may also be heaped on years with notable events (e.g., the year of independence, changes of government), which would not be captured by this index.

2.3 TRENDS IN THE MEDIAN AGE AT MARRIAGE

Age at first marriage was calculated as the difference between a woman's date of birth and her date of first union. Thus, this variable incorporates inaccuracies in both the date of birth as well the date of first union. Some insight into the quality of data can be obtained by calculating the median age at marriage by age group and comparing observed and expected patterns. These statistics for cohorts 20-24 through 45-49 years are shown in Table 2.3. When progressing from older to younger cohorts, the median age at marriage is expected either to increase or remain constant. A pattern of declining values or a U-shaped pattern would suggest problems with the data.

Table 2.3 Median age at first union, by age of woman at the time of the survey, Demographic and Health Surveys, 1986-1989

	Age of Woman at the Time of Survey					
Country	20-24	25-29	30-34	35-39	40-44	45-49
SUB-SAHARAN AFRICA						
Botswana	a	a	23.8	22.6	23.5	25.0
Burundi	a	19.5	19.5	19.6	19.4	19.7
Ghana	18.7	18.5	18.1	18.1	17.6	17.8
Liberia	18.2	17.9	17.2	17.2	16.0	16.6
Mali	15.9	15.9	15.6	15.6	15.6	15.8
Senegal	17.2	16.7	16.5	16.2	16.1	15.9
Togo	18.6	18.4	17.7	18.5	18.0	18.7
Uganda	17.8	17.5	17.0	16.8	16.6	16.7
Zimbabwe	19.7	18.8	18.5	19.0	18.1	18.6
NORTH AFRICA						
Morocco	a	19.9	18.9	18.1	16.8	16.3
Tunisia	a	22.9	21.3	20.6	19.4	19.8
ASIA						
Indonesia	19.6	18.1	17.6	16.8	16.4	16.5
Sri Lanka	a	23.2	22.7	23.1	21.4	20.0
Thailand	a	21.0	20.6	20.4	20.2	19.5
LATIN AMERICA/CARIBBEAN						
Brazil	a	21.1	21.5	21.2	20.7	NA
Colombia	a	20.8	21.3	21.1	20.0	20.2
Dominican Republic	19.9	19.3	18.6	18.2	17.7	18.2
Ecuador	a	19.8	20.2	20.2	19.9	20.5
Guatemala	18.9	18.5	18.5	18.6	19.1	NA
Mexico	a	20.2	19.4	20.5	19.7	19.2
Peru	a	21.0	20.7	20.3	20.3	19.3
Trinidad & Tobago	19.7	19.8	19.7	19.7	19.8	18.8

NA = Not available

[a] Fewer than half of the women in the age group have ever been married by the lower limit of the age group.

Values for median age at marriage behave as expected in most surveys for cohorts 20-24 through 40-44 years. For successively younger women, the median age either increases or remains stable in all three Asian surveys, in eight of the eleven African surveys, and in five of the eight surveys in Latin America and the Caribbean. Four of the surveys which do not conform to the expected pattern show erratic values—Botswana, Mexico, Togo, and Zimbabwe. In two other surveys (Brazil and Colombia), the pattern departs from expectation in that median age declines between age groups 30-34 and 25-29 years.

A comparison of median ages at marriage in the two oldest cohorts (women 40-44 and 45-49), deviates from the expected pattern in many surveys. The median age at marriage is at least one-half year older for the oldest cohort than for the next cohort in four African surveys (Botswana, Liberia, Togo, and Zimbabwe) and in three surveys in Latin America and the Caribbean (Dominican Republic, Ecuador, and Guatemala). The reason for this is most likely forward displacement of first unions in the oldest cohorts. In addition, in many African and Latin American countries, informal consensual unions are common and, in countries where this is the practice, the oldest women may fail to recall or to report accurately the date of their first informal union. A final anomaly is the case of Trinidad and Tobago where the median age at marriage is inexplicably low for women 45-49.

For 13 countries that participated in the World Fertility Survey (WFS), it is possible to compare median age at marriage in the WFS and DHS surveys for the same age cohorts. These comparisons are shown in Figure 2.1. In most of the countries, the medians match reasonably well. The median ages reported in the two surveys are usually within one-half year of each other and trends across cohorts are generally similar.

Figure 2.1 Median age at first union, selected WFS and DHS surveys, by age at the time of the DHS survey

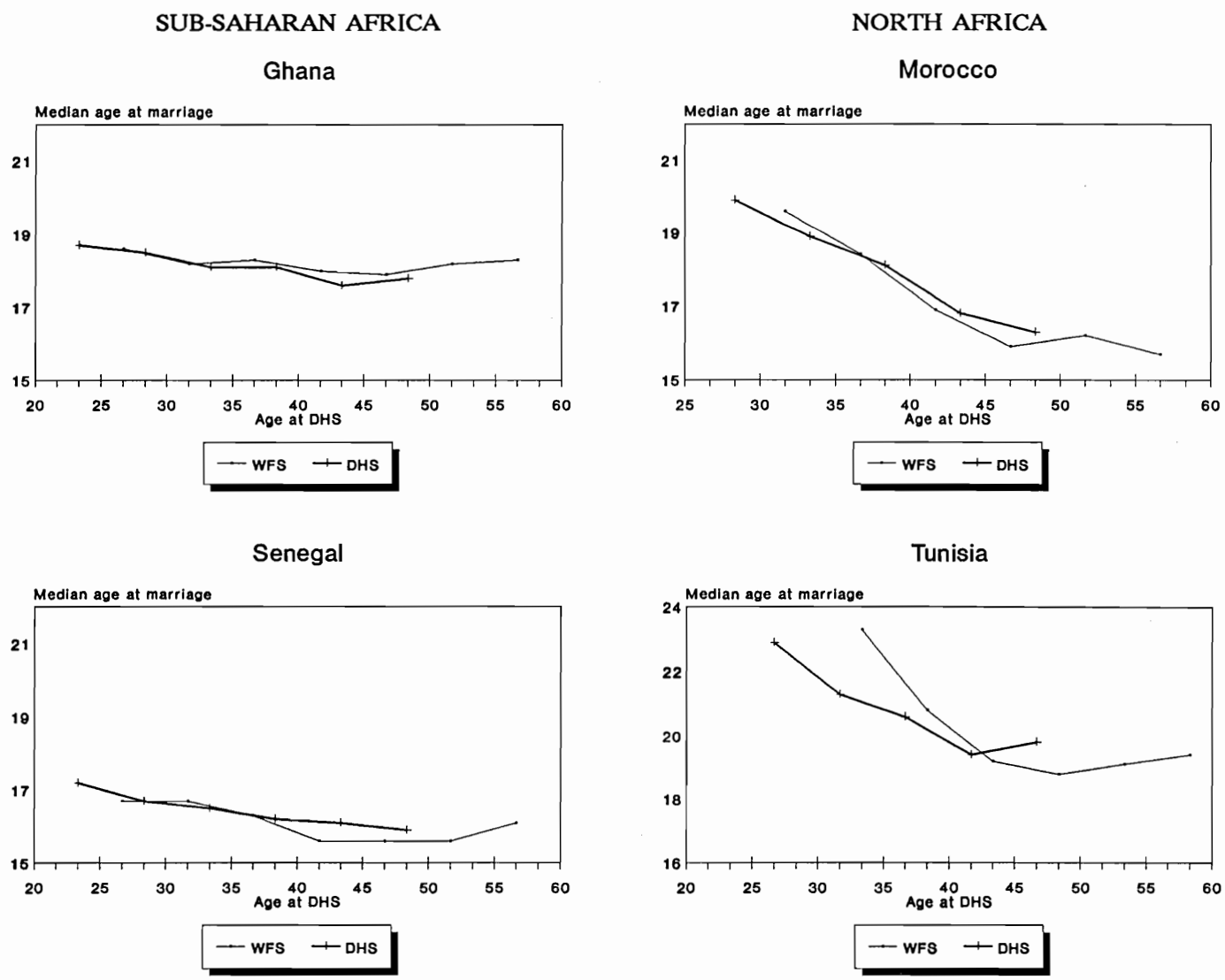
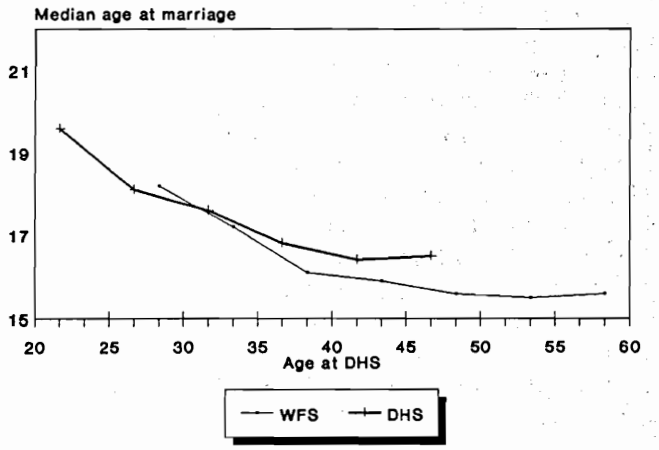


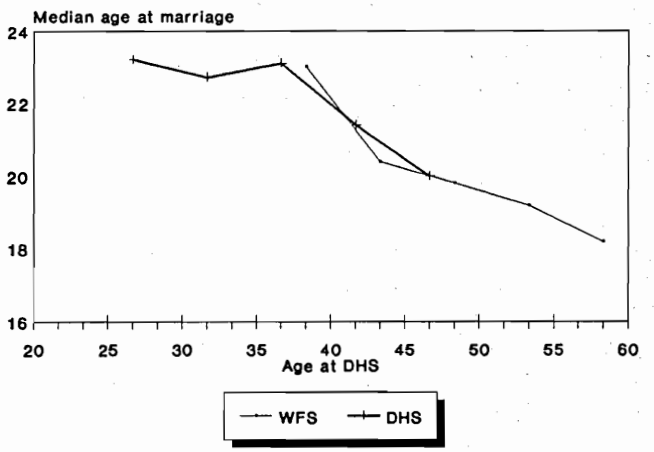
Figure 2.1—*Continued*

ASIA

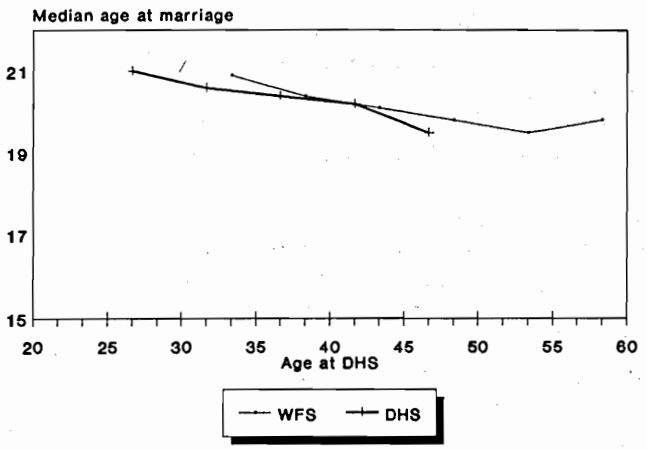
Indonesia



Sri Lanka

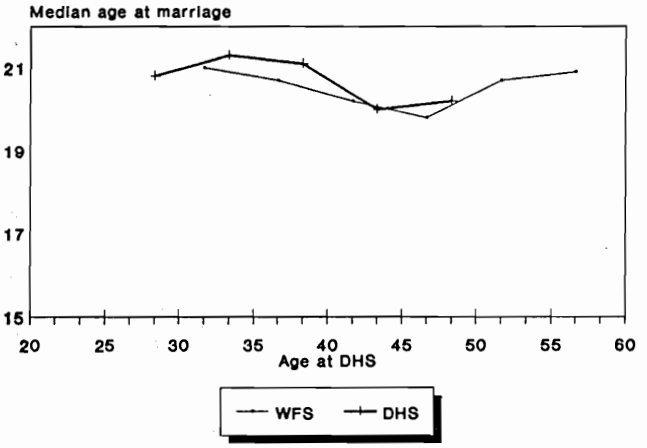


Thailand

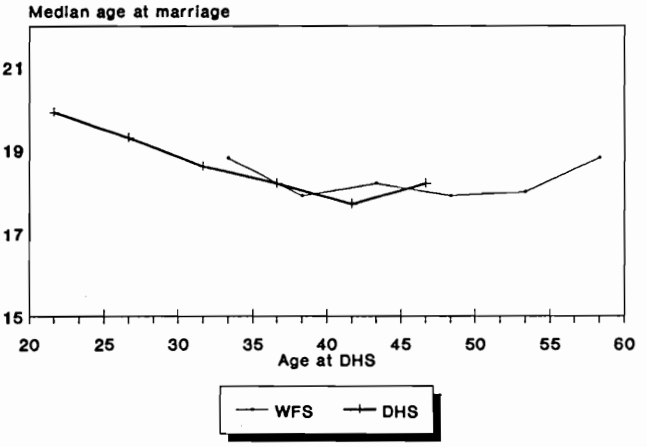


LATIN AMERICA/CARIBBEAN

Colombia



Dominican Republic



Ecuador

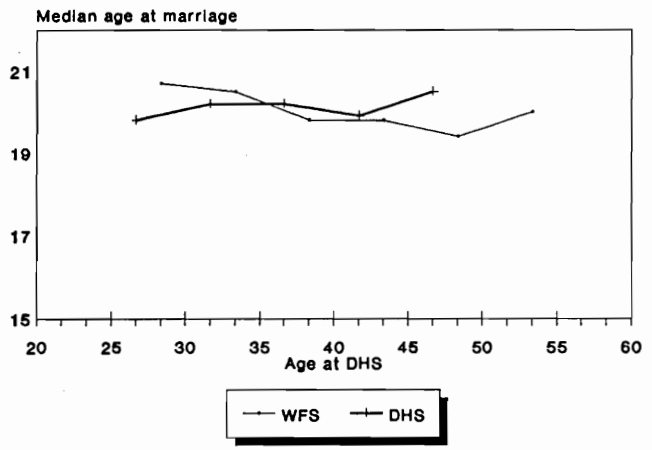
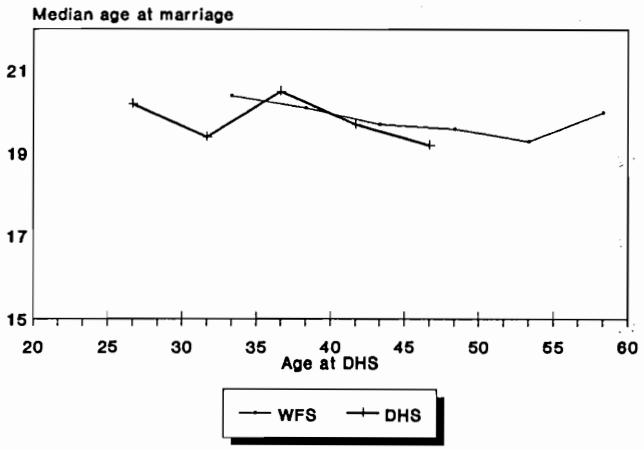


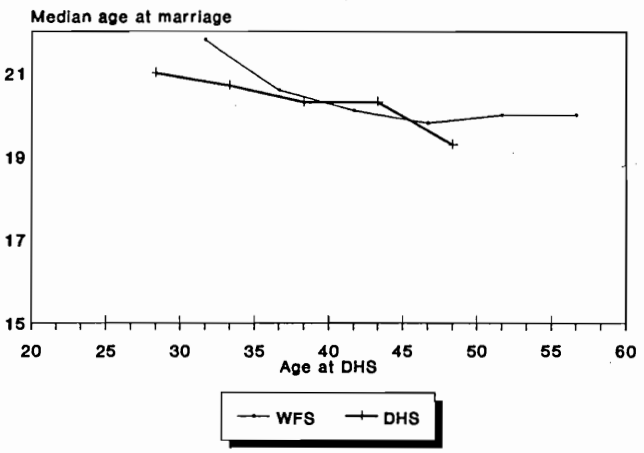
Figure 2.1—*Continued*

LATIN AMERICA/CARIBBEAN

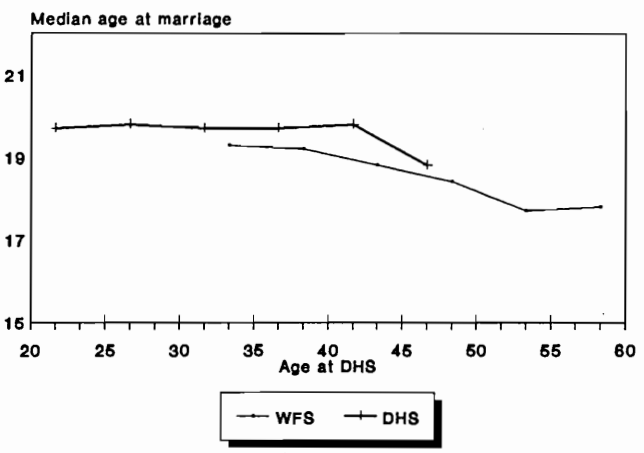
Mexico



Peru



Trinidad and Tobago



2.4 PROPORTION OF WOMEN EVER MARRIED AT AGES 15-19 AND 20-24

Since it is generally not possible to calculate the median age at marriage for women 15-19 at the time of the survey, and it is possible to make the calculation for women 20-24 in only half of the surveys, the quality of marriage data for these age groups will be examined by looking at the proportion of women ever married in cohorts 15-19 and 20-24 years (at the time of the survey) and exactly 5, 10, and 15 years prior to the survey (Table 2.4). The proportion of women ever married at the time of the survey is derived directly from the current marital status of women 15-19 and 20-24. The proportion ever married at ages 15-19 and 20-24 five years before the survey are calculated from the date of first union for women 20-24 and 25-29, respectively, at the time of the survey; the proportion ever married 10 years ago is calculated from the date of first union of women 30-34 and 35-39 at the time of the survey, and so on. In general, the proportion of women ever married in each age group is expected to decline or to remain stable, moving forward in time, and the proportion ever married at the time of the survey—derived from current status data—should be consistent with trends observed for the three previous periods as estimated from the retrospective data.

As expected, the proportion ever married among women 15-19 and 20-24 generally decreases from the distant to the recent past. There are a few cases, however, in which an increase is observed. The proportion of women ever married at age 15-19 ten years ago compared with 15 years ago increases in Botswana, Brazil, Ecuador, and Thailand. The proportion ever married at age 20-24 ten years ago compared with 15 years ago increases in Brazil, Burundi, Guatemala, Mexico, Sri Lanka, Togo, and Zimbabwe. An increase between 10 and 5 years prior to the survey in the proportion ever married among women 15-19 is found in Brazil, Guatemala, and Trinidad and Tobago, and among women 20-24 in Brazil, Colombia, Ecuador, Guatemala, and Trinidad and Tobago. In Mali, the proportion ever married among women 15-19 and 20-24 is higher at the time of the survey than for any preceding period.

The trend toward a decreasing proportion of women ever married in the youngest age groups accelerates in the most recent period for a number of countries. A doubling of the percentage decline in the proportion ever married among women 15-19 at the time of the survey, compared with years in the past, is seen in Brazil, Burundi, Colombia, Dominican Republic, Ghana, Guatemala, Indonesia, Sri Lanka, and Togo. This accelerated decline suggests that

Table 2.4 Percentage of women ever married at ages 15-19 and 20-24, for selected years prior to the survey, Demographic and Health Surveys, 1986-1989

Country	Percentage Ever Married Among Women 15-19, For Selected Years Before The Survey				Percentage Ever Married Among Women 20-24, For Selected Years Before The Survey			
	0	5	10	15	0	5	10	15
SUB-SAHARAN AFRICA								
Botswana	6.1	9.6	15.2	13.1	30.3	39.8	44.0	49.3
Burundi	6.8	15.1	20.0	20.4	66.7	74.7	74.4	72.5
Ghana	24.4	32.2	34.6	38.7	77.4	80.3	82.3	83.1
Liberia	36.0	40.3	44.9	47.0	75.3	79.2	78.8	81.6
Mali	75.4	66.5	67.2	71.3	98.0	94.5	96.0	96.5
Senegal	43.5	47.3	50.0	54.3	77.4	84.6	87.6	91.5
Togo	27.2	34.8	36.1	43.0	75.8	81.3	86.4	80.8
Uganda	40.8	44.1	44.1	51.2	83.0	85.0	85.7	88.0
Zimbabwe	19.8	26.7	34.6	34.5	71.6	79.8	82.5	78.2
NORTH AFRICA								
Morocco	12.3	18.8	27.4	35.1	44.7	61.7	70.9	77.1
Tunisia	4.4	7.9	10.2	13.0	35.9	45.4	57.6	61.2
ASIA								
Indonesia	19.0	33.7	42.6	48.2	65.2	77.8	80.3	84.5
Sri Lanka	6.9	11.8	12.5	12.7	42.3	45.1	47.9	46.1
Thailand	17.4	17.5	20.5	18.9	52.0	58.7	60.4	61.8
LATIN AMERICA/CARIBBEAN								
Brazil	14.7	18.3	17.9	16.9	55.8	59.0	57.2	56.4
Colombia	14.2	19.1	19.7	21.1	51.8	59.0	57.1	57.6
Dominican Republic	21.8	29.0	32.1	37.4	63.2	69.2	73.9	80.2
Ecuador	19.2	23.9	28.8	23.0	59.5	68.0	65.4	65.4
Guatemala	26.2	36.9	36.0	38.5	70.2	76.4	75.2	73.9
Mexico	19.9	24.4	29.5	32.1	58.3	65.0	68.4	62.4
Peru	12.8	17.4	20.4	22.4	48.8	58.7	63.8	64.2
Trinidad & Tobago	24.6	28.5	22.8	25.8	68.2	71.4	70.2	71.9

problems may exist in the reporting of marital status or age among young women.

Some studies suggest that there is a tendency for young married women to overstate their current age and for young unmarried women to understate their age (Coale, 1988; Makinson, 1984). If this occurs, the proportion ever married among young women will tend to be underestimated. This seems likely to be the problem in Burundi, Colombia, Guatemala, Indonesia, and Togo, since large declines in the current period persist when the proportion of women who ever had a first birth at age 15-19 are examined (see section 3.4). In Brazil, Dominican Republic, Ghana, and Sri Lanka, however, there are no large decreases in the proportion who had a birth in the current period, an indication that the problem in these four countries probably stems from misreporting of marital status or date of marriage, rather than misreporting of age. An exaggerated decline in the proportion ever married in the recent past will appear if newly married women tend to overestimate the duration of their marriage—that is, the proportion of women ever married five years prior to the survey will be overestimated while the current proportion of women ever married will be accurate (United Nations, 1987a). Another possibility is that young women underreport very early unions or current (probably consensual) unions. In the case of Sri Lanka, an ever-married sample, misreporting of the marital status of young women in the household survey may have contributed to an underestimation of the proportion of women ever married in this age group.

The proportion ever married at age 20-24 tends to show a smaller decline for the current period than the comparable proportion among women 15-19. An accelerating decline in the proportion of women ever married at age 20-24 can be seen in five countries in which there was a recent increase in the age at marriage—Indonesia, Mexico, Morocco, Peru, and Thailand. In six other countries, Burundi, Colombia, Ecuador, Guatemala, Senegal, and Zimbabwe, where there is no apparent trend toward increasing age at marriage, the decline in the proportion of women ever married may be a consequence of newly married women overestimating the duration of their marriage.

2.5 COMPARISON OF DHS SURVEYS AND OTHER SOURCES

Another test of the quality of reporting of union dates is the comparison of DHS data with data on marital status from previous censuses or surveys. Using information on the date of first marriage from a DHS survey, it is possible to reconstruct the retrospective proportion of women ever married at the time of a previous census or survey and compare this reconstruction with the proportion ever married reported in the earlier census or survey. It is expected that the retrospective estimates of the proportion of women ever married calculated from DHS data will be higher than the estimates from previous censuses or surveys for three reasons. First, since censuses often use a less inclusive definition of marriage than that of DHS surveys, the proportion of women ever married (as calculated from DHS data) can be expected to be higher than that estimated in a census. Apart from some countries in Latin America and the Caribbean, many censuses include as ever married only those women who were ever legally married and, thus, count as single (or never-married) women currently or previously in informal unions. Higher estimates of the proportion of ever-married women due to the inclusion of informal unions in the definition of marriage in DHS surveys should be particularly noticeable in younger age groups in which women who had early informal unions of short duration, but are not currently married, would be counted as ever-married in a DHS survey, but may be classified as single in a census.

Second, information on marital status and date of marriage in DHS surveys usually comes from the individual questionnaire, for which the respondent is a woman, rather than from the household questionnaire, for which the respondent is often a male head of household.[5] A third factor which might act to improve the validity of estimates from DHS surveys, relative to earlier censuses and surveys, is that the quality of reporting of birth and marriage dates may have improved in recent years, as the level of female education increased and countries placed greater emphasis on the recording of vital events.

The current proportion of women ever married by age from census or survey data may be affected by misreporting of marital status or current age. As noted previously, some studies have demonstrated a tendency for the age of young married women, especially those with children, to be overstated, and for the age of young unmarried women, especially those without children, to be understated. This type of age misreporting is expected to be more severe in censuses, where reporting is usually done by proxy, than in surveys, where the woman herself is interviewed.[6]

If this type of age misreporting occurs, the proportion of women ever married in the younger age groups will be too low, producing an exaggerated trend toward increasing age at marriage. Comparisons of proportions of women ever married from sources of data collected close together in time, would be expected to yield smaller discrepancies than comparisons between sources further apart in time. This is because recent comparisons are based on the reports of women of the same or similar ages and the magnitude of the age misreporting would be expected to be similar for both sources. Further back in time, the proportion of women ever married at young ages, reconstructed from DHS survey data, is increasingly based on the reports of older women who are presumed to report their current age more accurately; thus, moving back in time, the reconstructed proportion of women ever married for each age group may be more accurate than the proportion based on the age reported in the census or survey conducted at that time (Coale, 1988; Makinson, 1984). At the same time, as noted earlier, the reporting of marriage dates, especially by older women, is often distorted, and this problem will affect comparisons with earlier datasets. In addition, if younger women overestimate the duration of their marriage, then the estimate of the proportion ever married in a recent period prior to the survey will be too high.

Table 2.5 presents the proportion of women ever married by five-year age groups in one or more recent censuses or surveys and the same proportion reconstructed for the date of those sources using DHS data. Uganda is excluded because no recent data source could be found.

DHS data generally exhibit a higher proportion of women ever married in each age group than is observed for other data sources. The largest differences occur in cohorts 15-19 and 20-24 years. In a few cases, the differences are greater than 10 percentage points (e.g., women 20-24 in Botswana, and 15-19 in Guatemala and Mali). In Trinidad and Tobago, the proportion of women ever married in the DHS survey exceeds that recorded in the census by 17 percentage points for both age groups. The large differences in Trinidad and Tobago are probably due to the inclusion of visiting unions in the DHS survey. Of the 42

[5] For ever-married samples, the proportion ever married at the time of the survey is derived from the household questionnaire. When retrospective estimates are constructed, however, the proportion ever married in the past is increasingly derived from the date of union reported in the individual questionnaire.

[6] In surveys where a substantial number of women do not know their age or birth date, and interviewers estimate these on the questionnaire, the same type of age misreporting may occur.

Table 2.5 Comparison of the percentage of women ever married in DHS surveys with the percentage ever married in previous censuses or surveys, by age at the time of the previous census or survey

Country	Source	Date of Reconstruction	Percentage of Women Ever Married at Age: 15-19	20-24	25-29	30-34	35-39	40-44	45-49
SUB-SAHARAN AFRICA									
Botswana	DHS	1981	13.1	42.2	60.6	73.2	75.2	82.8	--
	Census	1981	7.3	31.2	53.1	67.6	74.8	78.9	82.6
Burundi	DHS	1979	21.7	76.3	94.6	98.1	98.7	98.7	--
	Census	1979	19.2	72.6	89.9	94.8	96.6	97.4	96.5
Ghana	DHS	1979/80	39.2	84.8	98.0	99.2	100.0	100.0	--
	WFS	1979/80	30.9	84.6	97.0	99.1	99.2	99.5	99.8
Liberia	DHS	1984	38.1	77.0	92.7	96.0	99.1	98.4	98.2
	Census	1984	35.7	70.9	86.1	92.7	95.5	96.4	96.9
	DHS	1974	45.0	81.1	93.7	96.7	95.8	--	--
	Census	1974	42.3	78.6	90.9	95.0	96.4	97.0	98.2
Mali	DHS	1976	66.2	95.9	99.4	98.9	100.0	--	--
	Census	1976	51.1	88.0	95.9	97.5	98.0	98.1	98.3
Senegal	DHS	1978	52.5	87.1	96.9	99.4	100.0	100.0	--
	WFS	1978	59.1	85.9	95.6	99.8	100.0	99.7	99.6
	DHS	1976	49.0	87.4	98.1	99.3	100.0	--	--
	Census	1976	38.6	76.1	92.3	96.9	97.3	97.5	97.6
Togo	DHS	1981	40.1	81.8	97.0	97.3	99.6	100.0	--
	Census	1981	43.3	81.8	94.0	96.6	97.4	97.2	96.8
Zimbabwe	DHS	1982	29.9	80.7	95.1	97.4	98.9	98.7	--
	Census	1982	26.1	76.5	90.7	94.5	96.4	97.0	97.2
NORTH AFRICA									
Morocco	DHS	1982	21.3	63.1	85.6	93.5	98.0	98.5	99.0
	Census	1982	18.5	59.6	83.0	93.6	97.9	99.0	99.2
	DHS	1979	26.8	68.6	87.9	97.0	98.2	99.0	--
	WFS	1979	21.3	63.1	86.6	95.7	98.9	98.7	99.2
Tunisia	DHS	1984	7.3	45.0	80.2	93.2	95.9	93.4	99.9
	Census	1984	6.7	41.0	75.4	90.3	96.2	97.8	98.4
	DHS	1978	9.8	56.7	83.2	93.1	93.1	--	--
	WFS	1978	5.4	43.7	80.1	94.0	98.1	98.3	98.8
ASIA									
Indonesia	DHS	1985	24.6	73.0	90.8	96.6	97.5	98.7	98.8
	SUPAS	1985	18.8	70.3	91.1	95.9	97.5	98.3	98.6
	DHS[a]	1976	46.5	82.5	94.5	98.0	98.8	--	--
	WFS[a]	1976	37.4	79.8	94.9	98.0	98.5	99.2	99.3
Sri Lanka	DHS[b]	1981	12.7	45.7	72.7	86.6	93.2	96.7	--
	Census	1981	9.9	44.7	69.6	84.2	91.1	94.1	95.5
	DHS[b]	1975	12.0	47.3	72.9	90.6	95.6	--	--
	WFS	1975	6.8	39.4	68.1	86.2	94.2	95.3	97.9
Thailand	DHS	1984	17.4	55.7	79.5	88.0	92.0	95.8	96.1
	SPC	1984	19.3	59.9	81.5	89.5	92.9	95.0	96.4
	DHS	1980	19.0	60.1	81.6	90.2	95.2	96.1	--
	Census	1980	16.7	56.5	79.1	88.2	92.7	94.7	95.9
LATIN AMERICA/CARIBBEAN									
Brazil	DHS	1980	18.8	59.7	82.3	91.2	93.5	--	--
	Census	1980	14.9	52.5	75.9	85.3	89.4	90.9	91.7
	DHS	1976	18.0	56.7	82.4	——91.4[c]——		——--[d]——	
	PNAD	1976	11.9	48.4	73.6	——87.0[c]——		——92.0[d]——	

Table 2.5—*Continued*

Country	Source	Date of Reconstruction	Percentage of Women Ever Married at Age: 15-19	20-24	25-29	30-34	35-39	40-44	45-49
Colombia	DHS	1980	20.6	59.0	79.4	87.3	92.7	95.6	--
	CPS	1980	13.6	51.1	74.9	82.7	88.5	91.9	93.5
	DHS	1977	20.0	57.8	78.5	89.2	92.6	96.1	--
	WFS	1977	15.1	56.0	77.7	88.6	87.7	91.2	90.9
Dominican Republic	DHS	1983	28.1	67.8	87.4	96.4	97.4	98.4	98.3
	CPS	1983	24.6	67.8	88.9	96.8	98.6	99.2	99.1
	DHS	1975	34.5	75.7	94.7	96.7	97.6	--	--
	WFS	1975	27.9	73.3	90.3	95.2	98.0	97.5	97.0
Ecuador	DHS	1982	22.7	66.8	85.5	90.8	93.2	97.5	100.0
	Census	1982	18.9	57.7	78.9	86.1	88.5	89.4	89.6
	DHS	1979	25.8	66.2	85.1	89.8	95.2	98.4	--
	WFS	1979	18.5	57.5	80.4	89.5	92.6	91.8	94.0
Guatemala	DHS	1987	28.5	71.7	90.4	94.1	95.6	96.8	--
	ENSD	1987	25.4	69.3	87.5	93.5	96.2	97.2	96.2
	DHS	1981	39.4	78.1	90.0	94.5	96.2	--	--
	Census	1981	27.9	69.3	85.2	90.7	92.9	93.6	94.2
Mexico	DHS	1980	26.8	64.4	85.0	90.9	93.3	94.8	--
	Census	1980	20.5	60.0	81.6	88.8	91.5	92.6	93.0
	DHS	1976	28.7	68.3	84.3	92.3	93.8	--	--
	WFS	1976	19.2	65.6	84.9	91.2	94.5	93.9	95.3
Peru	DHS	1981	18.9	58.9	83.1	93.6	95.7	95.6	--
	Census	1981	14.5	51.1	75.0	85.5	89.6	90.9	91.5
	DHS	1977	20.2	60.4	87.6	93.1	96.1	91.8	--
	WFS	1977	14.0	51.7	76.7	89.1	91.7	94.9	94.8
Trinidad & Tobago	DHS	1980	28.5	70.7	91.4	94.1	95.9	98.7	--
	Census	1980	11.5	53.8	80.6	89.2	91.9	93.5	94.3
	DHS	1977	22.1	70.8	87.7	94.3	98.2	--	--
	WFS	1977	39.0	68.4	90.1	95.1	96.7	98.1	97.8

CPS = Contraceptive Prevalence Surveys
ENSD = Encuesta Nacional Socio-Demográfica
PNAD = Pesquisa Nacional por Amostra de Domicilios
SPC = Survey of Population Change
SUPAS = Intercensal Population Survey
WFS = World Fertility Survey

[a]
[b] Java-Bali only
[c] Excludes approximately 14 percent of the estimated 1986 total population
[d] Women 30-39 years
Women 40-49 years

comparisons of the two youngest age groups in the 21 countries, the DHS estimate of the proportion of women ever married is equal to or less than 5 percentage points higher than the estimate from other data in twenty-one, between 5 and 10 percentage points higher in twelve, more than 10 percentage points higher in five, and less than the other data source in four. In six countries—Burundi, Liberia, Peru, and Sri Lanka, Tunisia, and Zimbabwe—the difference between the two sources is greater in age group 25-29 than in either of the two younger age groups.

There is little evidence of large regional variations in the magnitude of the differences between the two sources except that, even excluding Trinidad and Tobago, the dif-

ferences tend to be slightly larger in Latin America than in Africa or Asia. In addition, while the average size of the differential decreases across age groups in all three regions, it decreases less in Latin America, suggesting that the inclusion of informal unions in DHS surveys is a main factor in the higher proportion of women ever married reported in the surveys. A similar regional pattern was described by Singh (1985) for the World Fertility Survey (WFS) data.

For 15 countries, a comparison of DHS data with a second (earlier) data source is shown in Table 2.5. As expected, the discrepancies in the proportion of women ever married between the DHS survey and the earlier source tend to be greater than the differences between the DHS survey and the more recent source. This is generally true, even when the earlier source is a survey and the more recent source is a census. Two exceptions to this pattern are Liberia and Trinidad and Tobago. In Liberia, the differences between the DHS survey and the 1974 census are smaller in every age group except women 15-19, than the differences between the DHS survey and the 1984 census. In Trinidad and Tobago, the proportion of women ever married reported in the DHS survey is significantly larger than that reported in the 1980 census, but slightly lower (except in age group 15-19[7]) than the proportion reported in the 1977 WFS.

Overall, the proportion of women ever married reported in the DHS surveys is reasonably consistent with the proportions reported in other sources for women 25-49. A large number of discrepancies appear in the two youngest age groups and, in these cases, the DHS surveys usually show a higher proportion of women ever married. There are three likely reasons for the higher proportion of women ever married in DHS surveys. First, in many cases, especially when comparing DHS data with census data, the more inclusive DHS definition of marriage affects the number of women considered ever-married. Second, the current proportion of women ever married reported in other sources may be underestimated due to age misreporting among young women. Finally, retrospective estimates of the proportion of women ever married—constructed from DHS data—may be too large, as a result of misreporting of the date of union among young women and, in particular, overestimation of the duration of marriage.

[7] The DHS survey in Trinidad and Tobago included all women while the WFS survey excluded women who were full-time students. The WFS restriction would be expected to increase the proportion of women ever married, especially in the youngest age group; thus, when the DHS and WFS surveys are compared, the proportion of women ever married in age group 15-19 is significantly lower in the DHS survey than in the WFS survey.

3 Date of First Birth

Age at first birth has been suggested as a more appropriate indicator of the beginning of exposure to childbearing than age at first union, especially in countries where informal unions and premarital births are common. Information on a woman's childbearing history was gathered by two procedures in the DHS surveys. First, each respondent was asked about the number of sons and daughters she had living with her, sons and daughters living elsewhere, and sons and daughters who had died. Then she was asked to provide a full live birth history in which the name, date of birth, sex, survival status, and age at last birthday, or age at death, was collected for each birth.

The intent of these detailed questions is to prompt the woman to report all the births she has ever had and especially to remind her of births that occurred a long time ago, as well as births of children who may have died early in life. Age at first birth is calculated by subtracting the woman's birth date from the birth date of her first child. Thus, the accuracy of information on age at first birth depends on the accuracy of the respondents reporting of both her own birth date and the birth date of her first child.

When the date of birth of a woman's first child is reported inaccurately, it not only affects the calculation of her age at first birth but also is likely to affect the accuracy of the birth dates of subsequent children. Since the birth history in DHS surveys begins with the first child, its placement in time may affect the placement of following births, especially when a woman is unsure of the exact birth date of her children but is able to estimate the length of the birth interval.

3.1 COMPLETENESS OF INFORMATION

The proportion of respondents in each survey who reported the year and month of their first birth and the proportion of cases for which some information was imputed is shown in Table 3.1. As in the comparable table for date of first union, the columns are ordered from left to right according to the degree of imputation. Column one shows the percentage of cases for which no imputation was performed. Column two is the percentage of cases for which a year was reported but a month was imputed. Column three indicates the percentage of cases for which only the current age of

Table 3.1 Percent distribution of women who have ever given birth by completeness of information on date of first birth, by age group, Demographic and Health Surveys, 1986-1989

Country	Year and Month Reported, No Imputation	Year Reported, Month Imputed	Age Reported, Year and Month Imputed	No Information Year and Month Imputed	Total
SUB-SAHARAN AFRICA					
Botswana					
20-24	99.3	0.7	0.0	0.0	100.0
30-34	95.8	3.8	0.0	0.5	100.0
45-49	93.7	4.9	0.0	1.4	100.0
Total	97.0	2.3	0.2	0.5	100.0
Burundi					
20-24	96.8	3.2	0.0	0.0	100.0
30-34	78.1	21.9	0.0	0.0	100.0
45-49	58.8	40.7	0.0	0.5	100.0
Total	79.0	20.9	0.0	0.1	100.0
Ghana					
20-24	88.2	5.3	6.6	0.0	100.0
30-34	76.4	9.4	14.1	0.2	100.0
45-49	62.2	13.9	23.3	0.6	100.0
Total	77.6	9.7	12.6	0.1	100.0
Liberia					
20-24	90.6	9.1	0.0	0.4	100.0
30-34	82.6	15.1	0.3	1.9	100.0
45-49	70.1	22.2	3.6	4.1	100.0
Total	84.0	13.4	1.0	1.6	100.0

Table 3.1—*Continued*

Country	Year and Month Reported, No Imputation	Year Reported, Month Imputed	Age Reported, Year and Month Imputed	No Information Year and Month Imputed	Total
Mali					
20-24	41.8	29.5	20.8	8.0	100.0
30-34	27.8	35.0	24.4	12.8	100.0
45-49	24.4	36.8	20.7	18.0	100.0
Total	33.3	31.5	23.8	11.4	100.0
Senegal					
20-24	86.0	13.7	0.0	0.3	100.0
30-34	70.0	29.0	0.0	1.0	100.0
45-49	59.9	39.7	0.0	0.4	100.0
Total	73.7	25.4	0.0	0.8	100.0
Togo					
20-24	73.8	8.8	16.8	0.6	100.0
30-34	45.0	20.9	33.0	1.1	100.0
45-49	46.5	36.4	43.6	3.4	100.0
Total	51.3	19.1	28.4	1.1	100.0
Uganda					
20-24	99.9	0.0	0.0	0.1	100.0
30-34	100.0	0.0	0.0	0.0	100.0
45-49	100.0	0.0	0.0	0.0	100.0
Total	100.0	0.0	0.0	0.0	100.0
Zimbabwe					
20-24	99.8	0.2	0.0	0.0	100.0
30-34	99.5	0.4	0.0	0.2	100.0
45-49	98.9	0.7	0.0	0.4	100.0
Total	99.4	0.4	0.0	0.2	100.0
NORTH AFRICA					
Morocco[a]					
20-24	69.2 (22.9)	0.4	7.4	0.0	100.0
30-34	57.4 (30.3)	1.5	10.2	0.6	100.0
45-49	29.9 (43.5)	8.2	17.6	0.8	100.0
Total	53.5 (31.0)	3.2	11.6	0.6	100.0
Tunisia					
20-24	97.5	2.1	0.0	0.5	100.0
30-34	96.9	1.7	0.4	1.0	100.0
45-49	87.9	7.4	1.1	3.5	100.0
Total	94.6	3.1	0.6	1.7	100.0
ASIA					
Indonesia					
20-24	93.4	6.6	0.0	0.0	100.0
30-34	83.1	16.6	0.2	0.1	100.0
45-49	61.7	38.2	0.0	0.1	100.0
Total	80.6	19.4	0.0	0.0	100.0
Sri Lanka					
20-24	99.1	0.4	0.2	0.4	100.0
30-34	96.7	2.4	0.7	0.2	100.0
45-49	84.8	9.0	4.4	1.8	100.0
Total	94.9	3.2	1.3	0.6	100.0
Thailand					
20-24	96.5	1.5	0.6	1.3	100.0
30-34	94.4	2.8	1.2	1.7	100.0
45-49	82.5	8.7	2.5	6.3	100.0
Total	92.2	3.7	1.4	2.7	100.0

Table 3.1—*Continued*

Country	Year and Month Reported, No Imputation	Year Reported, Month Imputed	Age Reported, Year and Month Imputed	No Information Year and Month Imputed	Total
LATIN AMERICA/CARIBBEAN					
Brazil					
20-24	99.6	0.4	0.0	0.0	100.0
30-34	98.1	1.0	0.9	0.0	100.0
40-44	96.5	1.5	1.7	0.2	100.0
Total	98.3	0.9	0.7	0.1	100.0
Colombia					
20-24	99.7	0.3	0.0	0.0	100.0
30-34	99.0	1.0	0.0	0.0	100.0
45-49	97.1	2.9	0.0	0.0	100.0
Total	98.9	1.1	0.0	0.0	100.0
Dominican Republic					
20-24	99.9	0.1	0.0	0.0	100.0
30-34	99.2	0.8	0.0	0.0	100.0
45-49	97.8	2.2	0.0	0.0	100.0
Total	99.0	1.0	0.0	0.0	100.0
Ecuador					
20-24	98.4	1.4	0.2	0.0	100.0
30-34	96.7	2.5	0.6	0.2	100.0
45-49	89.5	7.2	2.2	1.1	100.0
Total	95.9	3.0	0.9	0.2	100.0
Guatemala					
20-24	99.0	0.9	0.0	0.1	100.0
30-34	95.9	3.6	0.0	0.5	100.0
40-44	92.6	6.5	0.0	0.8	100.0
Total	96.6	2.9	0.0	0.4	100.0
Mexico					
20-24	99.4	0.4	0.2	0.0	100.0
30-34	98.8	0.7	0.4	0.0	100.0
45-49	98.0	1.6	0.4	0.0	100.0
Total	98.8	0.9	0.2	0.0	100.0
Peru					
20-24	99.3	0.7	0.0	0.0	100.0
30-34	99.0	1.0	0.0	0.0	100.0
45-49	98.7	1.3	0.0	0.0	100.0
Total	99.1	0.8	0.0	0.0	100.0
Trinidad & Tobago					
20-24	100.0	0.0	0.0	0.0	100.0
30-34	99.8	0.2	0.0	0.0	100.0
45-49	97.0	2.7	0.0	0.4	100.0
Total	99.3	0.5	0.0	0.2	100.0

[a] Numbers in parentheses are percentage of women who reported a season and year.

the first born child was reported and the year and month of the first birth were imputed based on that age. Column four shows the percentage of cases for which no information was provided and the year and month were imputed based on surrounding events (e.g., age at first sexual intercourse and date of second birth).

A much larger proportion of women provided a year and month for their first birth than for their first union. In only seven countries, predominantly in Latin America, did more than 80 percent of ever-married respondents report a year and month of first union. However, in 16 of the 22 countries more than 80 percent of respondents with at least

one birth reported the year and month of their first birth and in 14 of those countries, more than 90 percent of the respondents gave a complete date. Furthermore, in those fourteen countries, even among the oldest women, 90 percent reported a year and month for their first birth. The date of first birth is obtained more frequently than the date of first union for a number of possible reasons: it can always be associated with a particular day while the marriage process can extend over several months or years, it is better documented (by birth certificates, school records, or immunization cards), and it is the subject of more probing during the interview.

More than 95 percent of the women in all the DHS surveys in Latin America and the Caribbean provided a complete date of first birth. Completeness of reporting drops off somewhat in the Asian countries, where 80 to 95 percent reported a complete date of first birth. In Africa, reporting of complete dates was high in Botswana, Morocco,[8] Tunisia, Uganda, and Zimbabwe, where nearly all women gave a complete date of first birth.

In West Africa, many more women reported a complete date of first birth than reported a date of first union. With the exception of Mali and Togo, three-quarters of the women in these countries provided a complete date of first birth. The poorest reporting was in Mali, where only one-third of the women reported both the year and month of their first birth. In Togo one-half of the respondents gave a complete date of first birth.

The proportion of birth dates not reported completely is greatest for older women. The decrease in completeness of reporting by age is especially large in the African datasets. For example, 97 percent of women 20-24 in Burundi reported a complete date of first birth compared with 59 percent of women 45-49. Similar declines for older women are observed in Ghana, Liberia, Senegal, and in one Asian country—Indonesia. Thus, age at first birth for older women is more often determined on the basis of imputed data than is the case for younger women.

Imputation of the month of first birth for more than 25 percent of cases occurred only in Mali and Senegal. It was necessary to impute the year and month of first birth—based on the current age of the child—in more than 10 percent of cases in Ghana, Mali, Morocco, and Togo. In Mali, a date of first birth was imputed in 11 percent of cases, based on no information at all. In all of the other countries, there was negligible imputation based on no information.

[8] In Morocco, 54 percent reported a year and month and 31 percent reported a season for the date of their first birth.

3.2 HEAPING

Birth dates are obtained by asking for both the year and month of birth and the current age of the child (if living). In addition to possible digit preference for year of birth, the estimation of birth dates in the field may be expected to result in heaping on years since first birth, i.e., the age of the child, ending in zero or five.

Indices measuring heaping of the date of first birth for the years 1960, 1965, 1970, 1975, and 1980 and heaping of duration since first birth on durations of 5, 10, 15, 20, and 25 years ago are presented in Table 3.2. A value of 1.0 indicates no heaping on the specified year or duration, while an index over 1.05 indicates significant heaping.

Table 3.2 Indices of heaping on year of first birth and duration since first birth ending in 0 or 5, Demographic and Health Surveys, 1986-1989

Country	Year of First Birth	Duration Since First Birth
SUB-SAHARAN AFRICA		
Botswana	1.03	0.98
Burundi	1.05	1.01
Ghana	1.01	1.11
Liberia	1.20	1.13
Mali	1.00	1.04
Senegal	1.05	1.04
Togo	1.05	1.15
Uganda	1.04	1.02
Zimbabwe	1.05	0.89
NORTH AFRICA		
Morocco	1.16	0.93
Tunisia	1.03	1.00
ASIA		
Indonesia	1.12	1.05
Sri Lanka	1.05	1.01
Thailand	0.98	1.06
LATIN AMERICA/CARIBBEAN		
Brazil	0.96	1.01
Colombia	1.00	1.03
Dominican Republic	1.03	1.02
Ecuador	1.05	1.01
Guatemala	0.92	1.08
Mexico	1.01	1.05
Peru	0.97	1.08
Trinidad & Tobago	1.05	0.98

For year of first birth:

$$\text{Index} = \sum_{i=1}^{5} \frac{x_i}{(x_{i-2}) + (x_{i-1}) + x_i + (x_{i+1}) + (x_{i+2})}$$

where $i_1 = 1960$, $i_2 = 1965$, $i_3 = 1970$, $i_4 = 1975$, $i_5 = 1980$,
x = number of women with first birth in year i.

Index for duration since first birth is identical except
$i_1 = 5$, $i_2 = 10$, $i_3 = 15$, $i_4 = 20$, $i_5 = 25$,
x = number of women whose first birth occurred i years ago.

There is little evidence of heaping on calendar years ending in zero or five except in Indonesia, Liberia, and Morocco where the indices are 1.12, 1.20, and 1.16, respectively. In no other country does the index exceed 1.05. The substantial heaping in Liberia may be due to a tendency of interviewers to push children out of the five-year reference period used in determining eligibility for the detailed health status questions in the individual questionnaire (see Arnold, this volume). Since the survey was conducted in 1986, information on health status should have been collected for all children born since January 1981. By making some children slightly older (i.e., born in 1980), the interviewer could avoid having to ask the long series of questions. In fact, heaping is most severe for the year 1980. Other surveys conducted in 1986, however, do not show excessive heaping or much evidence of birth displacement across the five-year boundary.

The value of the index of heaping for duration since first birth (or age of the oldest child) varies from 0.89 in Zimbabwe to 1.15 in Togo. Aside from Togo, two additional countries have values greater than 1.10—Ghana and Liberia. The indices for Guatemala and Peru suggest heaping to a lesser degree.

3.3 TRENDS IN THE MEDIAN AGE AT FIRST BIRTH

Table 3.3 shows the median age at first birth for cohorts 20-24 years through 45-49 years. In the absence of a change in age at first birth, we would expect no change in these values across cohorts. Recent increases in age at first birth should be reflected in higher median ages for the younger cohorts. A trend toward lower median age at first birth is generally a result of the forward displacement of births by older women, i.e., closer to the survey date than actually occurred, in combination with accurate reporting by younger women. The omission of early births by older women would result in a similar pattern.

Table 3.3 Median age at first birth by age at the time of the survey, Demographic and Health Surveys, 1986-1989

	Age at the Time of Survey					
Country	20-24	25-29	30-34	35-39	40-44	45-49
SUB-SAHARAN AFRICA						
Botswana	19.7	19.2	19.3	19.6	20.0	20.9
Burundi	a	20.9	21.1	21.1	21.2	21.1
Ghana	19.9	20.0	19.2	19.5	18.8	19.3
Liberia	18.5	19.0	19.4	19.8	18.6	21.0
Mali	18.4	19.0	18.6	19.1	18.7	20.3
Senegal	19.0	19.0	19.0	18.7	18.7	19.2
Togo	19.5	19.2	18.8	19.5	19.2	20.0
Uganda	18.6	18.3	18.0	18.0	18.3	18.6
Zimbabwe	a	19.5	19.4	19.8	19.2	19.7
NORTH AFRICA						
Morocco	a	22.4	21.3	21.0	20.1	20.0
Tunisia	a	24.5	23.2	22.5	21.5	22.3
ASIA						
Indonesia	a	20.2	19.9	19.6	19.3	19.8
Sri Lanka	a	24.7	24.2	24.9	23.1	21.9
Thailand	a	23.0	22.7	22.3	22.2	21.6
LATIN AMERICA/CARIBBEAN						
Brazil	a	22.4	22.8	22.4	22.2	NA
Colombia	a	21.6	21.9	21.6	20.8	21.0
Dominican Republic	a	21.2	20.5	19.9	19.9	19.7
Ecuador	a	20.7	21.0	21.2	21.1	21.6
Guatemala	20.0	19.7	19.7	20.1	20.2	NA
Mexico	a	21.1	20.6	21.5	21.0	20.7
Peru	a	21.4	21.4	21.2	20.8	20.5
Trinidad & Tobago	a	22.2	21.9	21.6	21.1	20.5

NA = Not available

[a] Fewer than half of the women in the age group have given birth by the lower limit of the age group.

The median age at first birth increases steadily from age group 40-44 years to 20-24 years in the Dominican Republic, Indonesia, Morocco, Peru, Senegal, Thailand, Trinidad and Tobago, and Tunisia. In addition, a recent increase in age at first birth is indicated in Colombia, Ghana, Sri Lanka, and Uganda. In Ecuador and Liberia, the trend is toward decreasing age at first birth for younger cohorts. It is unclear whether the latter is a real trend or whether younger women have displaced their first birth backwards and older women have displaced it forwards.

Half of the surveys show a greater age at first birth for the oldest cohort (45-49 years) than for the next oldest cohort (or cohorts). For example, in Liberia, women 45-49 years had a median age at first birth more than two years older than women 40-44 years; the difference in Mali is more than one and one-half years. The only surveys in Africa which do not show evidence of omission or forward displacement of births by women in the oldest cohort are Burundi and Morocco. Displacement of births by older women is less severe in Asia and Latin America, occurring only in Colombia, Ecuador, and Indonesia.

Overall, the data presented in Table 3.3 indicate that in a number of DHS countries, particularly in sub-Saharan Africa, the date of the first birth for older women is being shifted closer to the time of the interview because of recall error or omission of earlier births. Displacement is generally restricted to women 40-49, however, and thus predominantly affects the estimation of trends in age at first birth and fertility for periods 15 or more years before the survey.

Figure 3.1 presents a comparison of median age at first birth in the DHS and WFS surveys in 13 countries (for the same age cohorts). The medians for age at first birth match

Figure 3.1 Median age at first birth, selected WFS and DHS surveys, by age at the time of the DHS survey

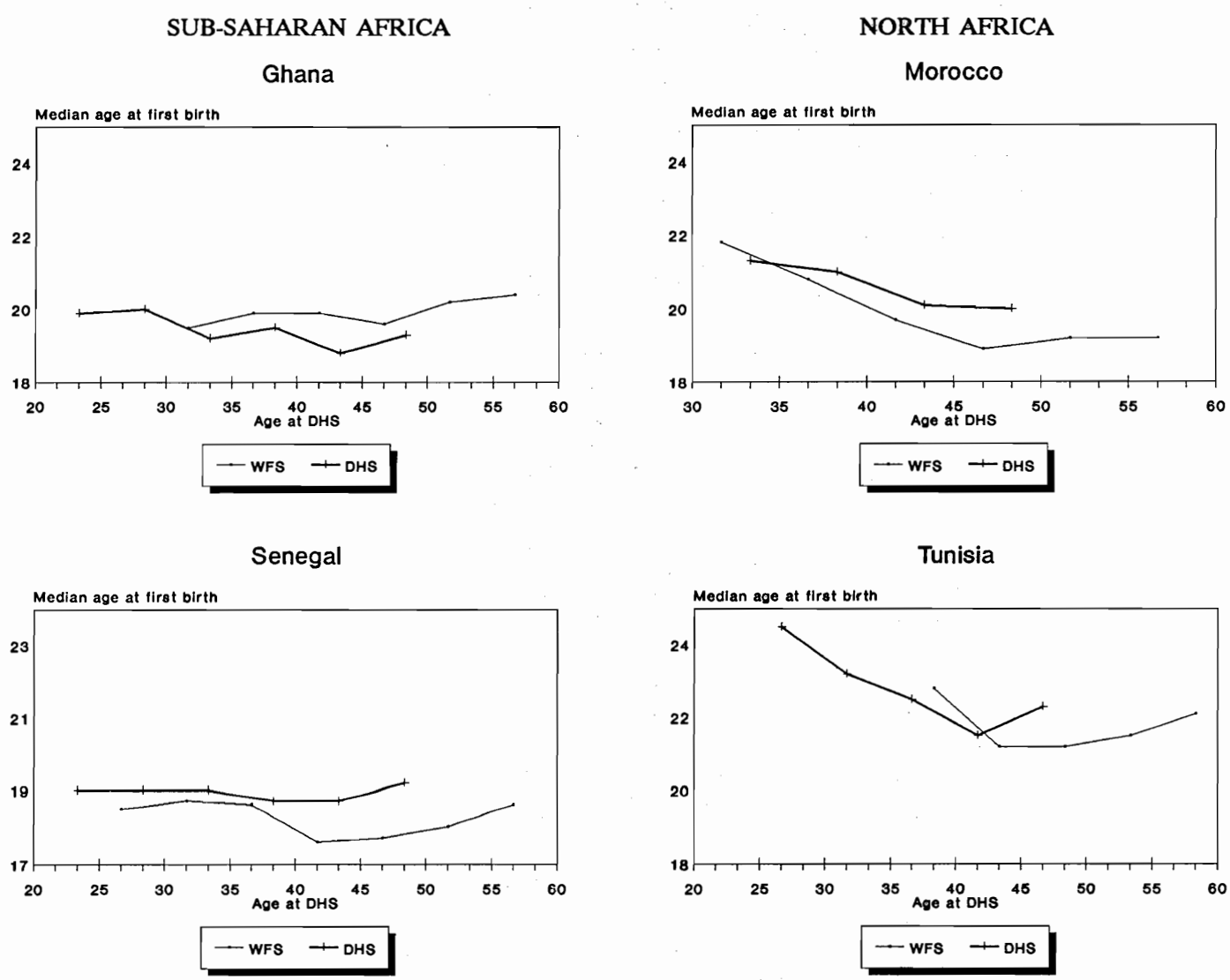
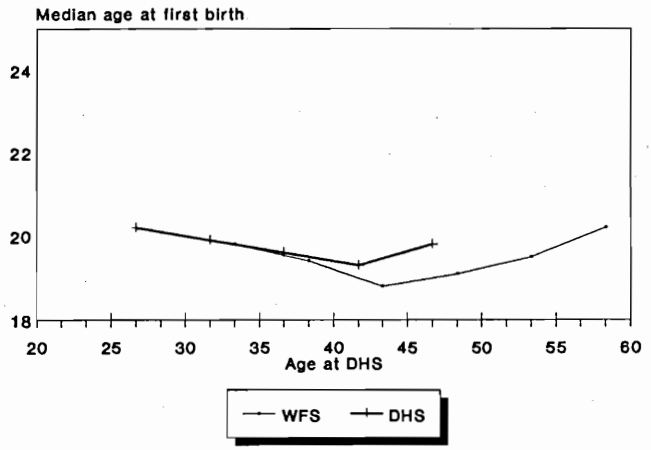


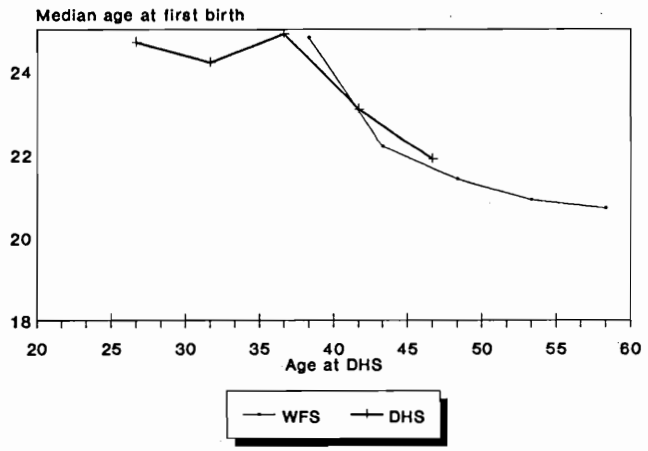
Figure 3.1—*Continued*

ASIA

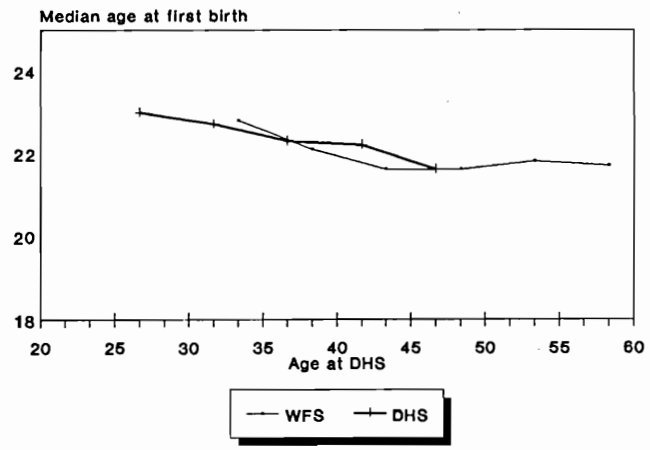
Indonesia



Sri Lanka

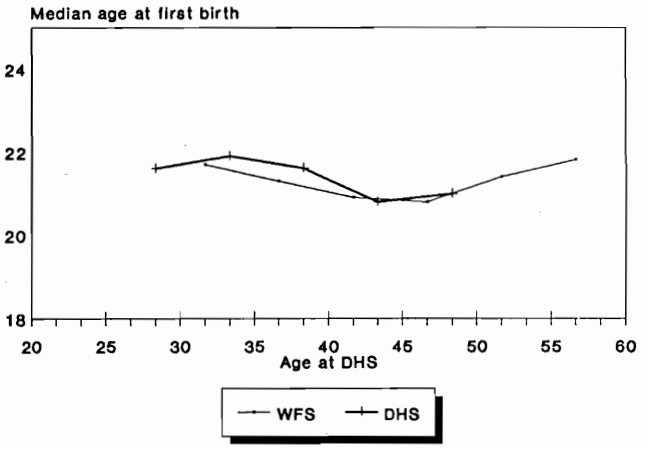


Thailand

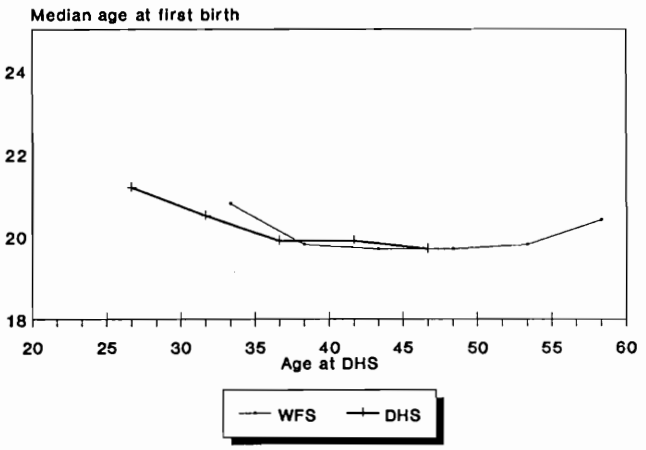


LATIN AMERICA/CARIBBEAN

Colombia



Dominican Republic



Ecuador

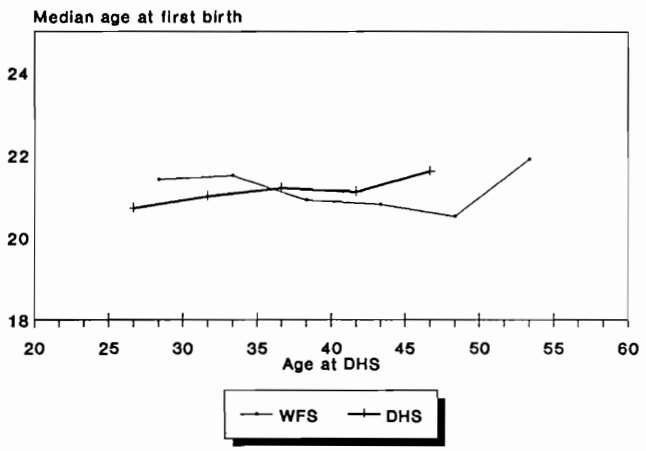
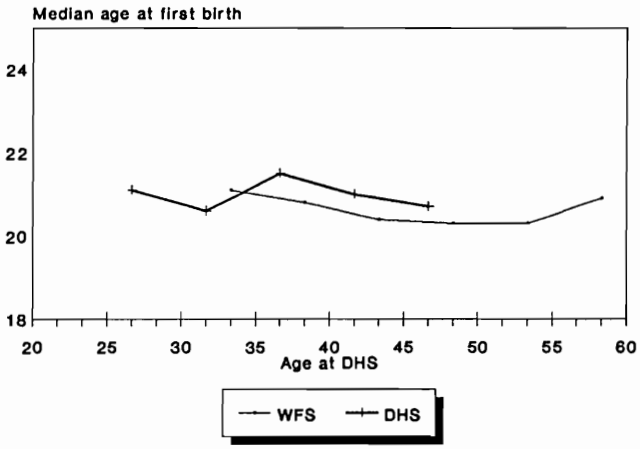


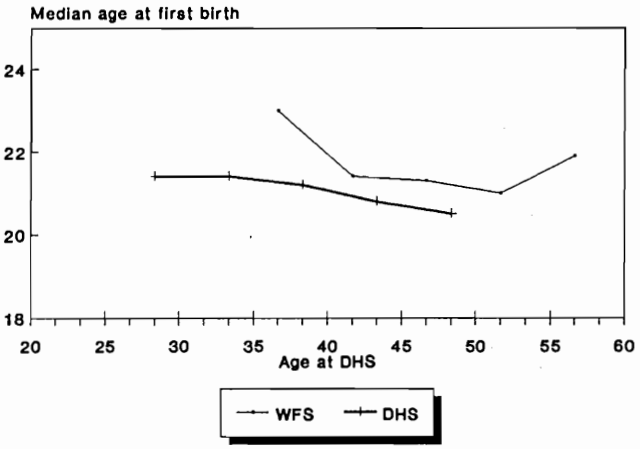
Figure 3.1—*Continued*

LATIN AMERICA/CARIBBEAN

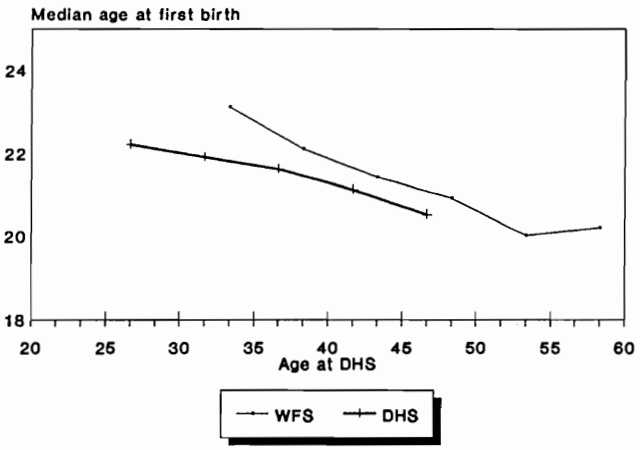
Mexico



Peru



Trinidad and Tobago



match less well, overall, than the medians for age at first marriage (shown in Figure 2.1). In Indonesia, Morocco, and Senegal, the two curves are similar for younger women, becoming less consistent for older cohorts. In Ecuador, the surveys show opposite trends while in Peru and Trinidad and Tobago, there is a gap of more than one year between the youngest comparable age cohorts. In Colombia, the Dominican Republic, Sri Lanka, and Thailand the match between curves is very close.

3.4 PROPORTION OF WOMEN WHO HAD A FIRST BIRTH AT AGES 15-19 AND 20-24

As in the case of age at first marriage, it is generally not possible to obtain a median age at first birth for women under 25. The quality of data on first births for young women is examined in Table 3.4, which shows the proportion of women 15-19 and 20-24 who ever had a birth at the time of the survey and 5, 10, and 15 years prior to the survey. The proportions would be expected to decrease or to remain stable over time. A trend towards increases in recent years—which would indicate earlier childbearing among young women—is questionable.

The proportion of women 15-19 who ever had a birth decreases steadily (or with minor deviation) in 6 of the 11 African surveys, all of the Asian surveys, and 4 of the 8 surveys in Latin America and the Caribbean. Among the remaining countries, there is no change in Brazil and Trinidad and Tobago, the pattern is erratic in Botswana, Colombia, Guatemala, Senegal, and Togo, and in Liberia and Mali, the proportion of women 15-19 who ever had a first birth increases from the distant to the recent past. The latter pattern could result from an increasing omission of early births among successively older cohorts of women, or from the displacement of early birth dates closer to the date of the interview.

In a number of countries, there is apparent omission or displacement of first births among women 20-24, 15 years before the survey. This may occur because data on births to women 20-24 are obtained from slightly older women—those 35-39 at the time of the survey—than the data on first births at age 15-19 (fifteen years ago), which was collected from women 30-34 at the time of the survey. The proportion of women who had a birth at age 20-24 fifteen years ago is smaller than that 10 years ago in every sub-Saharan country except Senegal and Uganda. The proportion also decreases among older women in Guatemala, Mexico, and Sri Lanka.

Table 3.4 Percentage of women who ever had a birth at ages 15-19 and 20-24, for selected years prior to the survey, Demographic and Health Surveys, 1986-1989

Country	Percentage Who Ever Had a First Birth Among Women 15-19, For Selected Years Before the Survey				Percentage Who Ever Had a First Birth Among Women 20-24, For Selected Years Before the Survey			
	0	5	10	15	0	5	10	15
SUB-SAHARAN AFRICA								
Botswana	23.5	22.0	26.3	19.7	74.6	81.5	79.7	76.9
Burundi	3.2	7.7	11.2	9.2	55.2	62.2	60.2	58.6
Ghana	19.3	19.7	20.7	23.9	72.1	70.5	74.7	71.9
Liberia	37.2	35.3	31.5	27.8	80.7	78.1	68.5	65.2
Mali	44.5	38.3	30.4	33.8	82.4	78.4	79.6	72.7
Senegal	26.1	29.8	28.4	29.0	73.6	75.9	75.9	76.0
Togo	21.4	27.2	24.7	29.4	74.0	75.3	80.0	72.4
Uganda	30.3	33.9	35.8	37.7	83.3	81.1	82.5	84.2
Zimbabwe	16.3	21.3	25.5	25.1	71.2	78.1	79.0	72.4
NORTH AFRICA								
Morocco	5.0	8.6	11.3	15.1	35.1	48.1	55.2	57.7
Tunisia	2.1	3.6	3.5	5.3	28.7	33.1	43.9	47.1
ASIA								
Indonesia	9.0	17.8	22.7	25.0	55.4	65.7	67.2	68.6
Sri Lanka	3.5	4.9	6.1	6.3	33.2	36.5	37.5	35.3
Thailand	7.8	9.0	9.2	8.9	40.2	46.3	47.9	49.6
LATIN AMERICA/CARIBBEAN								
Brazil	10.4	11.0	11.0	9.3	50.7	48.5	45.2	47.9
Colombia	10.5	14.3	13.5	14.1	49.8	54.8	53.1	55.8
Dominican Republic	13.9	15.6	19.2	20.5	53.7	57.0	61.3	66.2
Ecuador	14.0	16.0	17.4	15.2	54.1	61.5	58.9	59.1
Guatemala	19.7	24.5	23.6	24.7	66.6	71.4	68.9	68.3
Mexico	14.0	16.7	18.4	20.4	53.0	58.4	62.4	57.2
Peru	10.8	12.3	14.8	15.5	47.0	56.2	57.5	57.5
Trinidad & Tobago	11.0	12.9	12.5	11.8	46.4	49.8	51.9	53.7

3.5 INTERVAL BETWEEN DATE OF FIRST UNION AND DATE OF FIRST BIRTH

The percent distribution of the interval between first union and first birth (first birth interval) is shown in Table 3.5. Column one presents the proportion of first birth intervals which were negative, i.e., the first birth occurred prior to the union. A large proportion of negative first birth intervals may imply that first unions of brief duration were omitted or that dates of first union were displaced forward in time. Although it is also possible that some births were displaced backward in time, the evidence so far on displacement of births suggests that first births tend to be reported as occurring later than the true date, rather than earlier.

In countries where short, informal unions are common, large numbers of premarital births should probably not be viewed as an indication of misreporting.

The percentage of first births occurring prior to the first union ranges from zero in Morocco to 51 percent in Botswana. However, in each of the regions, the percentage of premarital first births is less wide-ranging and is generally consistent with accepted social practices: 12-24 percent in sub-Saharan Africa (Botswana and Burundi being the exceptions), under 3 percent in North Africa and Asia, and 4-25 percent in Latin America and the Caribbean.

In Botswana, teenage pregnancies are fairly common and the percentage of premarital births is more than twice as high as any other country. The high level of premarital pregnancies is probably related to the fact that there is little pressure on women to enter a union because of pregnancy. Similar social practices may also explain the high level of premarital births in Liberia and Togo (relative to the other sub-Saharan surveys).

Column four of Table 3.5 contains the percentage of first birth intervals which were longer than three years. Even in countries with moderate levels of contraceptive use, it is expected that the majority of first births will occur within three years of the beginning of the union. This would be especially true for countries with low levels of contraceptive use or little use of contraception for spacing early births. Thus, long first birth intervals may indicate that misreporting of the date of first birth or first union has occurred.

Table 3.5 Percent distribution of the interval between first union and first birth, Demographic and Health Surveys, 1986-1989

Country	Negative Interval	Months Between First Union and First Birth			Total
		0-7	8-35	36+	
SUB-SAHARAN AFRICA					
Botswana	50.9	7.9	33.3	8.0	100.0
Burundi	4.0	13.2	73.6	9.2	100.0
Ghana	11.8	14.4	62.0	11.8	100.0
Liberia	24.1	12.7	38.6	24.6	100.0
Mali	14.0	10.8	37.5	37.7	100.0
Senegal	14.5	13.0	48.7	23.8	100.0
Togo	21.9	17.9	46.3	13.9	100.0
Uganda	14.4	11.9	64.7	9.1	100.0
Zimbabwe	16.9	17.7	57.3	8.1	100.0
NORTH AFRICA					
Morocco	0.0	3.4	71.7	24.9	100.0
Tunisia	2.3	9.8	74.0	13.9	100.0
ASIA					
Indonesia	1.3	5.6	70.0	23.1	100.0
Sri Lanka	2.3	6.4	81.4	9.9	100.0
Thailand	0.3	1.9	85.0	12.9	100.0
LATIN AMERICA/CARIBBEAN					
Brazil	8.8	14.7	66.5	10.7	100.0
Colombia	11.8	12.1	68.8	7.4	100.0
Dominican Rep.	4.1	9.9	70.9	15.2	100.0
Ecuador	11.3	12.7	66.9	9.2	100.0
Guatemala	6.3	12.0	73.6	8.1	100.0
Mexico	8.2	13.4	69.5	8.8	100.0
Peru	14.6	19.3	59.9	6.2	100.0
Trinidad & Tob.	6.4	10.6	64.7	18.3	100.0

A large proportion of long first birth intervals (i.e., more than 15 percent of the intervals are longer than three years) occur in the Dominican Republic, Indonesia, Liberia, Mali, Morocco, Senegal, and Trinidad and Tobago. In Liberia, Mali, and Senegal this could be due to subfecundity in the first few years of marriage, due to the young age of women at marriage. A relatively high level of contraceptive use among women in union with no births is consistent with long first birth intervals in Trinidad and Tobago. For the remaining countries, the Dominican Republic, Morocco, and Trinidad and Tobago, the large proportion of long first birth intervals suggests the possible omission of early births, as well as forward displacement of births.

4 Age at First Sexual Intercourse

All ever-married women were asked the age when they first had sexual intercourse. Additionally, never-married women were asked if they ever had sexual intercourse; those replying affirmatively were also asked the age when they first had sexual relations. The question on age at first sexual intercourse was eliminated from the questionnaire in four countries: Morocco, Tunisia, Senegal, and Sri Lanka.

4.1 COMPLETENESS OF INFORMATION

As only age and not date of first intercourse was collected,[9] there is no comparable analysis on the completeness of reporting of year and month, as was presented for first union and first birth. Instead, the analysis looks at the percentage of cases for which the data were missing or for which a special code was present indicating that the respondent refused to answer the question (on age at first intercourse) or responded "don't know." Despite concerns of some researchers and survey staff that questions on sexual intercourse would make respondents uncomfortable and would be difficult for them to answer, it was found that among women interviewed in DHS surveys, who had had sexual intercourse, only a few did not provide their age at their first sexual intercourse.

The highest levels of non-response were found in Brazil and Mali; however, the percentage of cases without data in these two countries is only 2.4 and 2.7 percent, respectively. In four other surveys, between 1 and 2 percent of the women did not respond, and in the remaining 12 surveys in which the question was included, less than 1 percent of women did not answer the question. While this does not address the validity of the responses, it does show that in general women were willing and able to answer the question.

4.2 TRENDS IN THE MEDIAN AGE AT FIRST SEXUAL INTERCOURSE

The expected trend in the median age at first sexual intercourse is difficult to predict. In settings where sexual intercourse occurs, or is expected to occur, exclusively within marriage, the trend in the median age at first sexual intercourse should approximate the trend in age at first marriage. In countries where informal unions or early temporary unions prior to the first stable union are common, however, the trend in age at first sexual intercourse may be different from the trend in age at first marriage.

In the African countries, the trend in age at first sexual intercourse across cohorts tends either to be flat or to show a slight increase (Table 4.1). The one exception is Zimbabwe, which shows a fairly consistent trend toward increased age at first sexual intercourse from the oldest to the youngest cohorts. For three countries with data indicating an increase in age at marriage—Ghana, Liberia, and Uganda—the trend in age at first sexual intercourse across cohorts is much flatter than the trend for age at first union. For example, in Liberia, the median age at first union increases 2.2 years between women 40-44 and women 20-24 at the time of the survey, but the increase in age at first sexual intercourse was only 0.3 years.

In the countries of Latin America and the Caribbean, the pattern is somewhat irregular. Only the Dominican Republic and Peru show a substantial increase in age at first sexual intercourse, and both countries also show a significant increase in age at marriage. In the two Asian countries (Indonesia and Thailand), a trend toward increased age at first sex parallels increased age at marriage.

An abrupt rise in the median age at first sexual intercourse among the oldest women does not occur to any appreciable extent, in contrast to the findings for age at first union and first birth. Perhaps, the memory women have of their first sexual intercourse is sharper than either the memory of when they first began living with a partner or when they had their first birth (though this seems unlikely). In most societies, established norms exist regarding the appropriate age at which women may begin to have sexual relations. It is impossible to know to what degree this norm determines actual behaviour, though it is likely that there is a considerable amount of correlation.

4.3 INTERVAL BETWEEN AGE AT FIRST SEXUAL INTERCOURSE AND FIRST UNION

The average gap across all age groups between the median age at first sexual intercourse and the median age at first

[9] In two countries, Brazil and Guatemala, women under 25 were also asked the month and year they first had sexual intercourse.

Table 4.1 Median age at first sexual intercourse, by age of woman at the time of the survey, Demographic and Health Surveys, 1986-1989

Country	Age of Woman at the Time of Survey					
	20-24	25-29	30-34	35-39	40-44	45-49
SUB-SAHARAN AFRICA						
Botswana	17.3	17.2	17.3	17.2	17.3	18.1
Burundi	a	19.2	19.3	19.3	19.4	18.8
Ghana	16.7	16.6	16.3	16.5	16.4	16.5
Liberia	15.5	15.5	15.4	15.4	15.2	15.4
Mali	15.9	15.8	15.6	15.6	15.6	15.7
Togo	16.5	16.3	16.0	16.7	16.4	17.1
Uganda	15.9	15.6	15.6	15.8	15.4	15.3
Zimbabwe	18.3	17.5	17.3	17.4	16.7	16.9
ASIA						
Indonesia	19.6	18.1	17.7	17.0	16.7	16.8
Thailand	a	21.0	20.6	20.5	20.4	19.7
LATIN AMERICA/CARIBBEAN						
Brazil	a	19.9	20.6	20.4	19.9	NA
Colombia	a	19.6	19.7	19.4	19.0	18.7
Dominican Republic	19.5	18.9	18.2	17.8	17.5	17.7
Ecuador	19.9	18.6	19.1	19.0	19.1	19.1
Guatemala	18.4	18.1	17.9	17.8	18.4	NA
Mexico	a	19.8	19.0	20.0	19.5	18.9
Peru	a	19.3	19.2	18.9	18.7	18.3
Trinidad & Tobago	19.3	19.5	19.2	19.1	18.8	18.2

NA = Not available
[a] Fewer than half of the women in the age group have had sexual intercourse by the lower limit of the age group.

marriage ranges from zero in Mali to 6.3 years in Botswana. In the African countries where the question on age at first sexual intercourse was asked, the gap is generally between 1 and 2 years, except in Burundi where it is 0.3 years. The interval is close to zero in Indonesia and Thailand, and between 1 and 1.5 years in the Latin America and Caribbean region. Given differential patterns of union formation in the three regions, these differences appear to be reasonable. However, this consistency does not always hold at the individual level.

Table 4.2 shows the distribution of women by the length of the interval between age at first sexual intercourse and age at first union. The large proportion of women reporting that they had sex prior to their first union may indicate that reliance on the date of first union to capture all regular sexual exposure is insufficient. A further purpose of examining this distribution is to detect an indicator of the internal consistency of the data—the percentage reporting an age at first sexual intercourse subsequent to age at first union.

In the African countries, the percentage of women reporting an age at first sexual intercourse one or more years prior to their age at first union ranges from 12 percent in Mali to 67 percent in Botswana. The figure for Mali is suspiciously low. As seen in Table 3.5, 14 percent of all births occurred and an additional 11 percent were conceived prior to the date of the first union, yet only 12 percent of women reported premarital sex. This inconsistency suggests considerable inaccuracy in the reporting of either age at first sexual intercourse or age at first union.

The percentage of cases where age at first sexual intercourse exceeds the age at first union by one or more years is between 10 and 14 percent in every African country, except Burundi (19 percent).[10] Burundi also has a much lower proportion of premarital births than the other sub-Saharan African countries. Some explanations for this phenomenon are that age at first union was misreported as occurring at too young an age, reported age was converted incorrectly to a date in the field, or women tended to overestimate their age at first sexual intercourse.

In Indonesia and Thailand, more than 85 percent said they first had sex in the same year that their first union began. Very few reported having had sex before their first union,

[10] In some countries (e.g., Mali and Tunisia), it is common to formalize a marriage contract prior to the couple living together. Although respondents were asked to report the date they started living with their first husband or partner, some women may have reported the date the marriage was contracted rather than the date cohabitation with the husband began. In most of the countries in which this type of misreporting would be a factor, however, the question on age at first sex was not asked.

Table 4.2 Percentage distribution of the interval between first sexual intercourse and first union, Demographic and Health Surveys, 1986-1989

Country	Age at FSI Less Than Age at First Union by 4+ Yrs.	Age at FSI Less Than Age at First Union by 1-3 Yrs.	Age at FSI Same as Age at First Union	Age at FSI Greater Than Age at First Union by Exactly 1 Yr.	Age at FSI Greater Than Age at First Union by More Than 1 Yr.	Cannot Be Calculated[a]	Don't Know/ Refused/ Missing	Total
SUB-SAHARAN AFRICA								
Botswana	38.8	28.4	14.6	8.4	3.3	6.4	1.2	100.0
Burundi	3.3	16.1	60.3	18.1	1.0	1.0	0.1	100.0
Ghana	16.7	42.5	26.1	9.0	3.8	1.8	0.1	100.0
Liberia	24.6	33.5	28.3	5.9	4.4	0.0	3.4	100.0
Mali	3.0	9.0	66.0	7.1	4.4	7.7	2.8	100.0
Togo	18.0	46.2	24.2	8.6	0.9	1.3	0.8	100.0
Uganda	16.2	36.9	32.8	13.1	0.8	0.0	0.2	100.0
Zimbabwe	12.6	34.3	40.3	9.3	1.1	2.1	0.3	100.0
ASIA								
Indonesia	0.2	2.5	88.6	3.2	4.5	0.0	1.0	100.0
Thailand	0.6	3.1	86.7	6.2	2.8	0.0	0.5	100.0
LATIN AMERICA/CARIBBEAN								
Brazil	8.0	22.6	57.8	4.4	0.8	4.3	2.3	100.0
Colombia	10.0	30.4	49.9	8.7	1.0	0.1	0.0	100.0
Dominican Rep.	2.6	15.5	77.1	4.4	0.3	0.1	0.0	100.0
Ecuador	7.4	21.0	68.9	1.0	0.3	1.3	0.0	100.0
Guatemala	4.7	20.4	57.4	10.6	0.8	5.0	1.1	100.0
Mexico	4.4	15.8	54.6	16.6	7.4	0.0	1.2	100.0
Peru	10.8	34.0	52.1	2.8	0.2	0.2	0.0	100.0
Trinidad and Tob.	6.5	17.6	57.7	14.8	1.5	1.3	0.6	100.0

FSI = First sexual intercourse

[a] In some countries, if the age at first sexual intercourse was more than one year greater than the age at first birth, age at first sexual intercourse was changed to code 97 (inconsistent data). Where the code 97 appears, it is impossible to calculate the interval between age at first sexual intercourse and age at first union.

while 8 to 9 percent gave an age at first sexual intercourse subsequent to their age at first union.

In Latin America and the Caribbean, 50 percent or more of women reported that they first had sex in the same year they began their first union. The percentage reporting that they had sex for the first time one or more years before their first union ranges from 18 percent in Dominican Republic to 45 in Peru.

The percentage who reported an age at first sexual intercourse a year or more after their age at first union is substantial in some of the countries in Latin America and the Caribbean, reaching 11 percent in Guatemala, 16 percent in Trinidad and Tobago, and 24 percent in Mexico. The magnitude of the inconsistency is greatest in Mexico where 7 percent of the women gave an age at first sexual intercourse two or more years subsequent to their age at first union. Mexico is the only country which did not ask women to report their age at first union in cases where they could not provide the year. As a result, the date of first union may have been estimated by interviewers (and respondents) more frequently than in other countries, which may, in turn, have generated a greater number of inconsistencies.

The relatively high percentages of women reporting that they first had sex a year or more after the start of their first union indicates that either age at first sexual intercourse or date of (or age at) first union were frequently reported inaccurately. The fact that the majority of the inconsistent responses fall within one year of being consistent suggests that the inconsistency may arise when a woman reports that she first had sex when she got married and the interviewer either incorrectly estimates the respondent's age at first sexual intercourse from her date of marriage (e.g., by subtracting the year of marriage from the interview year without taking account of the respondent's month of birth or the month of the interview) or incorrectly works out the date of first union based on the age provided by the respondent. In fact, in 13 of the 18 countries included in Table 4.3, the percentage of respondents whose age at first sexual intercourse is a year or more subsequent to their age at first union is greater for women who reported a year of marriage than for those who reported an age.

Table 4.3 Percent distribution of women who have ever given birth by the interval between first sexual intercourse and first birth, Demographic and Health Surveys, 1986-1989

Country	Age at FSI Greater Than Age at First Birth by More Than One Year	Age at FSI Same as or 1 Yr. Greater Than Age at First Birth	Age at FSI Less Than Age at First Birth by 1-3 Yrs.	Age at FSI Less Than Age at First Birth by 4+ Yrs.	Don't Know/ Refused/ Missing	Total
SUB-SAHARAN AFRICA						
Botswana	6.1	15.3[a]	57.0	20.6	1.0	100.0
Burundi	1.1	18.5	69.5	10.7	0.1	100.0
Ghana	1.9	8.9a	60.8	28.4	0.1	100.0
Liberia	0.1	13.1	43.0	40.2	3.6	100.0
Mali	8.5	14.7[a]	41.7	32.1	2.9	100.0
Togo	1.5	15.9	50.9	31.0	0.7	100.0
Uganda	0.0	12.4	64.7	22.6	0.2	100.0
Zimbabwe	2.2	14.8	65.8	16.9	0.4	100.0
ASIA						
Indonesia	0.0	9.1	74.7	15.0	1.1	100.0
Thailand	0.6	13.9	75.5	9.4	0.5	100.0
LATIN AMERICA/CARIBBEAN						
Brazil	4.8	8.6[a]	69.5	14.7	2.4	100.0
Colombia	0.5	14.3	70.8	14.3	0.0	100.0
Dominican Rep.	0.1	9.8[a]	76.6	13.4	0.0	100.0
Ecuador	1.4	11.4[a]	73.6	13.5	0.0	100.0
Guatemala	5.5	11.3[a]	71.6	10.6	1.1	100.0
Mexico	2.2	26.8	61.8	8.1	1.2	100.0
Peru	0.2	11.4[a]	75.0	13.3	0.0	100.0
Trinidad and Tob.	1.6	11.2[a]	67.0	19.7	0.5	100.0

FSI = First sexual intercourse

[a] The "rule of one" was applied: if the age at first sexual intercourse was greater than the age at first birth by one year, the age at first sexual intercourse was decreased by one year, and there are no cases where the age at first sexual intercourse is exactly one year greater than age at first birth.

4.4 INTERVAL BETWEEN AGE AT FIRST SEXUAL INTERCOURSE AND FIRST BIRTH

The percent distribution of the length of intervals between age at first sexual intercourse and age at first birth is shown in Table 4.3. In most countries, age at first birth was checked against age at first sexual intercourse during secondary editing. If the age at first sexual intercourse was found to be greater than the age at first birth by one year, the age at first sexual intercourse was decreased by one year to make it consistent; thus, it is impossible to detect the extent to which this type of error occurred. If the age at first sexual intercourse was found to be greater than the age at first birth by more than one year, age at first sexual intercourse was changed to code 97 (inconsistent data).[11] These latter cases appear in column one of Table 4.3. The percentage of cases in which the respondent reported an age at first sexual intercourse more than one year greater than her age at first birth ranges from zero in Indonesia and Uganda to 8.5 percent in Mali. Between 3 and 6 percent of cases have this type of inconsistency in Botswana, Brazil, and Guatemala. In the remaining countries, 2 percent or fewer of the respondents reported an age at first sexual intercourse two or more years subsequent to their age at first birth.[12]

[11] In most surveys, age at first sex was used as a constraint on the imputation of both date of first union and date of first birth in cases where no other information was available. Thus, the number of cases imputed on the basis of no information (see Table 2.1 and Table 3.1) will affect the proportion of inconsistent responses.

[12] The proportion of cases in which the age at first birth was four or more years greater than the age at first sexual intercourse is also shown in Table 4.3. As previously mentioned, in countries with low levels of contraceptive use or little use of contraception for spacing early births, long first birth intervals may indicate misreporting. The detection of misreporting in these data is difficult, however, because the countries in which contraceptive use is very low (mainly in sub-Saharan Africa) are also those in which the interval between first sexual intercourse and first birth (i.e., the first sex-first birth interval) is long, as a result of subfecundity following early initiation of sexual relations. In fact, there is a strong negative association in these data between age at first sexual intercourse and the percentage of first sex-first birth intervals of more than three years (R^2 = .70).

5 Summary and Conclusions

The intent of collecting information on age at first union, age at first birth, and age at first sexual intercourse is to identify the beginning of the reproductive period in women's lives. In this report, we have examined the quality of this information in DHS surveys. The techniques used to detect deficiencies in the data include internal consistency checks, the evaluation of indicators of completeness of reporting and of heaping, and comparison of DHS data with data from other sources. The analysis has revealed some obvious deficiencies in the data, but it should be noted that the structure of this comparative analysis precludes detailed country-specific evaluation and comparison of DHS data with all available data sources. A summary of the data quality indicators by country is presented in Table 5.1.

An evaluation of the completeness of reporting of the date of first union shows that the extent to which respondents reported a month and year for this event varies between regions, with the most complete reporting in Latin America and the least complete in West Africa. The completeness of reporting also decreases with increasing age of the respondent. Nevertheless, in all of the surveys at least three-quarters of the women provided either a date of or age at first marriage.

Heaping on calendar years and durations since first marriage ending in zero or five was also examined. About half of the countries show some evidence of heaping on calendar years of first union. In general, more countries show heaping on calendar years than on durations since first marriage ending in zero or five and this problem is most apparent in the surveys in Latin America. However, Liberia shows evidence of significant heaping on both calendar years and durations of first marriage.

The magnitude of misreporting of the date of first union was investigated by looking at the patterns of median age at marriage across age cohorts. As expected, for age groups 20-24 through 40-44, the median age at marriage either decreases or remains stable in most surveys. There is evidence, however, of an appreciable amount of forward displacement of the date of first union in the oldest age cohort in most of the surveys. There is also some evidence of misreporting of age or marital status among the youngest women (primarily women 15-19). In several surveys, the current proportion of women ever married in this age group is low enough to suspect that age and marital status interact to cause the misclassification of a substantial proportion of young women. In particular, if the ages of young married women are overstated and the ages of young unmarried women are understated, the proportion of women ever married in the youngest age groups will be underestimated. There are indications that this has occurred in Burundi, Colombia, Guatemala, Indonesia, and Togo.

Comparison of the proportion of women ever married in DHS surveys and other sources, by age, generally shows a higher proportion married at all ages in the DHS surveys, but especially in the younger age cohorts. The data suggest that DHS surveys have obtained levels of coverage of first unions that are at least as complete, and in some cases more complete, than other sources. Nevertheless, the large discrepancies at ages 15-19 and 20-24 between DHS data and other sources in some countries are consistent with the underestimation of the proportion of women ever married in these age groups in other sources (a problem evident in DHS surveys as well, as noted above), and misporting of marriage dates among recently married women in the DHS surveys.

The evaluation of the quality of data on age at first birth parallels the evaluation of age at first marriage. A greater proportion of women were able to give a complete date of first birth than were able to give a complete date of first marriage. In 16 of 22 countries, more than 80 percent of respondents supplied a month and year of first birth. Heaping of dates of first birth on calendar years or durations ending in zero or five is less problematic than for dates of first union. Indonesia and Morocco show evidence of significant heaping on preferred calendar years while in Ghana and Togo there is noticeable heaping on duration since first birth. In Liberia, heaping on both calendar years and duration since first birth is severe.

In eight surveys, the median age at first birth increases steadily from age group 40-44 years to 20-24 years and a recent increase in age at first birth is indicated in four other surveys. In two countries, the trend is toward decreasing age at first birth for younger cohorts, suggesting some backward displacement of first births by younger women and forward displacement by older women. In more than half of the surveys, an increase in the median age at first birth among the oldest women indicates that forward displacement of the date of first birth, or omission of early births, has occurred. Omission or displacement of early births can also be seen in a comparison of the proportion of women who had a birth at age 20-24, at the time of the survey and 5, 10, and 15 years prior to the survey. This problem is most evident in sub-Saharan Africa.

Figure 5.1 shows a comparison of the medians for age at first sexual intercourse, age at first union, and age at first

Table 5.1 Summary of data quality indicators, Demographic and Health Surveys, 1986-1989

Data Quality Indicators	BOTSWANA	BURUNDI	GHANA	LIBERIA	MALI	SENEGAL	TOGO	UGANDA	ZIMBABWE	MOROCCO	TUNISIA	INDONESIA	SRI LANKA	THAILAND	BRAZIL	COLOMBIA	DOMINICAN REP	ECUADOR	GUATEMALA	MEXICO	PERU	TRINIDAD & T
FIRST MARRIAGE																						
Years and/or Month Imputed >30 %	X	X	X	X	X	X	X				X											
Index of Heaping on Year >1.05		X	X	X				X			X	X			X	X	X		X		X	X
Index of Heaping on Duration >1.05				X	X		X											X	X	X		
Median Age at Marriage for Women 45-49 >40-44 by 0.5+ Years	X			X			X		X						NA		X	X	NA			
Recent Accelerated Decline in Proportion Married at Age 15-19[a]		X	X				X	X				X	X		X	X	X		X			I
Erratic Trend in Median Age at Marriage for Age Groups 20-24 Through 40-44 Years[b]							X									X				X		
FIRST BIRTH																						
Year and/or Month Imputed >30%	X				X		X		X			X										
Index of Heaping on Year >1.05				X						X												
Index of Heaping on Duration >1.05			X	X			X							X							X	
Median Age at First Birth for Women 45-49 >40-44 by 0.5+ Years	X		X	X	X	X	X		X		X	X			NA			X	NA			
Recent Accelerated Decline in Proportion Who Ever Had a Birth at age 15-19[a]						I	I	X			X	X		X	X	I			I			X
Erratic Trend in Median Age at First Birth for Age Groups 20-24 Through 40-44 Years[b]				X	X								X							X		
AGE AT FIRST SEXUAL INTERCOURSE																						
Age at First Sex is Greater Than Age at First Union Among More Than 10% of Ever-Married Women	X	X	X	X	X	NA		X	X	NA	NA		NA						X	X		X
Age at First Sex is 1+ Years Greater Than Age at First Birth Among More Than 5% of Women Who Ever Had a First Birth	X				X	NA				NA	NA		NA						X			

I = Indeterminate
NA = Not Applicable

[a]Percentage decline from 5 years prior to the survey to the time of the survey is at least double the percentage change from 10 to 5 years prior to the survey.
[b]Pattern of two or more changes, 0.5 or more years, in opposite directions.

Figure 5.1 Median age at first sexual intercourse, first union, and first birth, Demographic and Health Surveys, 1986-1989

SUB-SAHARAN AFRICA

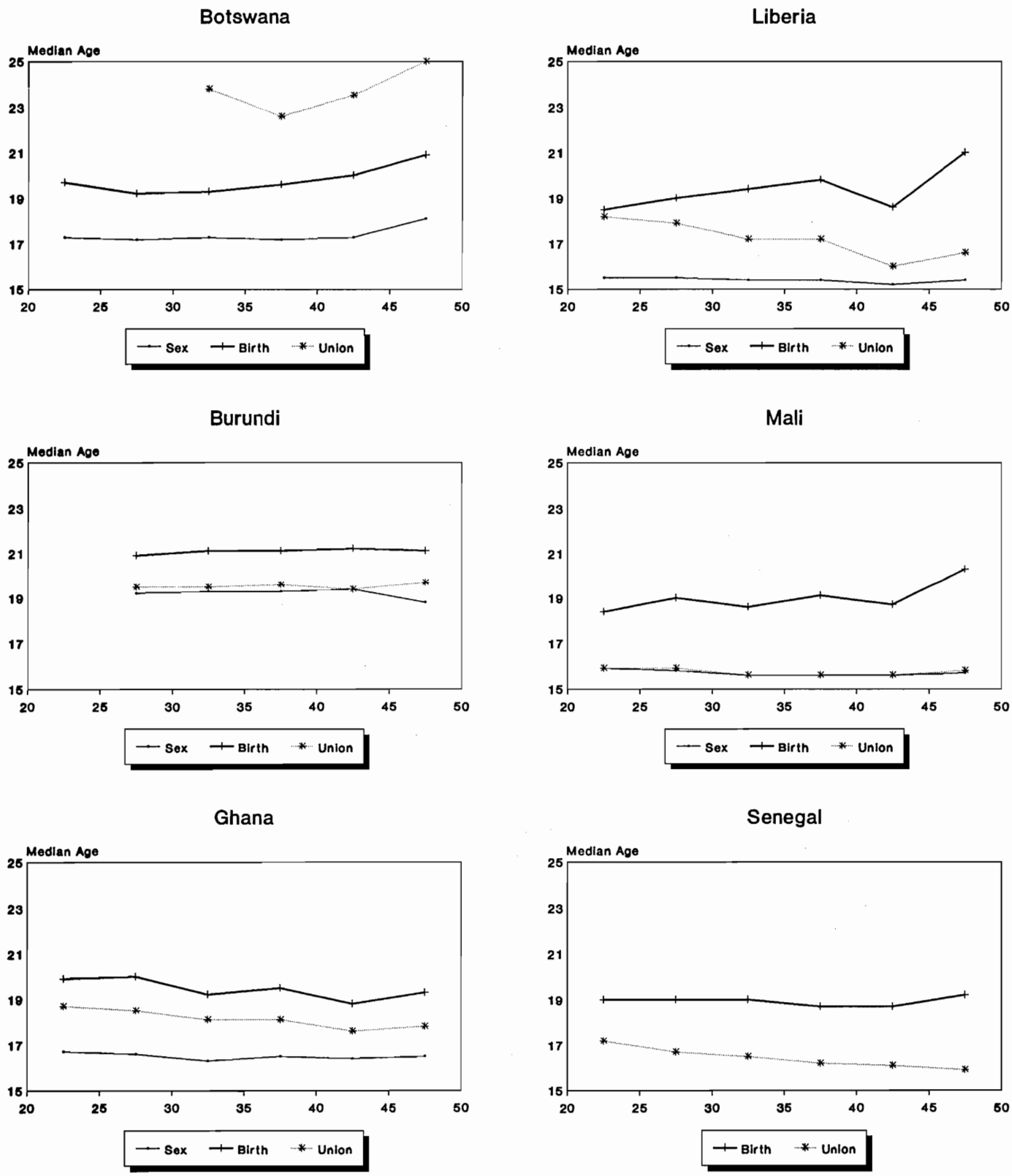
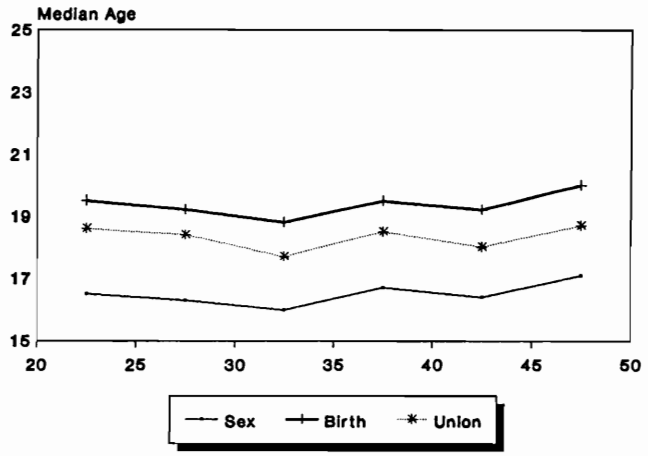


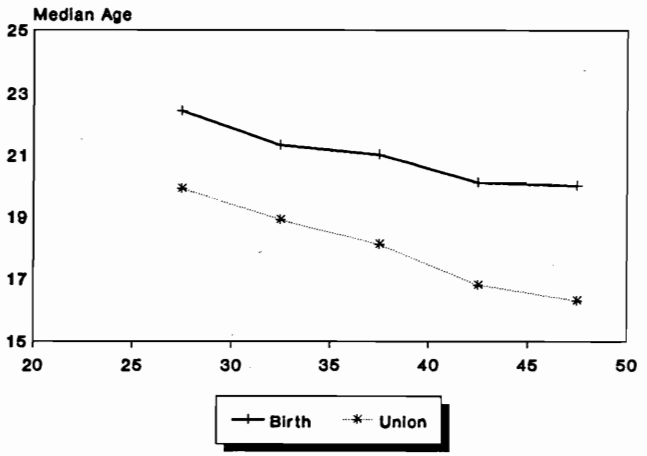
Figure 5.1—*Continued*

NORTH AFRICA

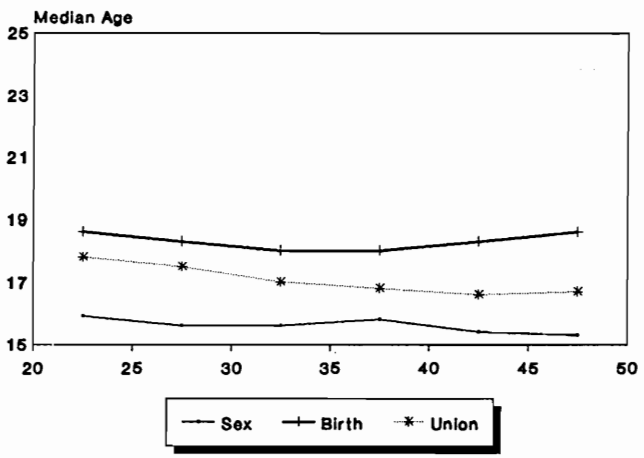
Togo



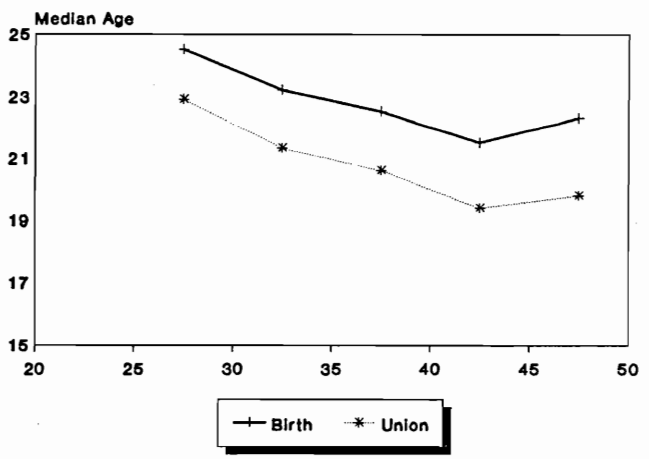
Morocco



Uganda



Tunisia



Zimbabwe

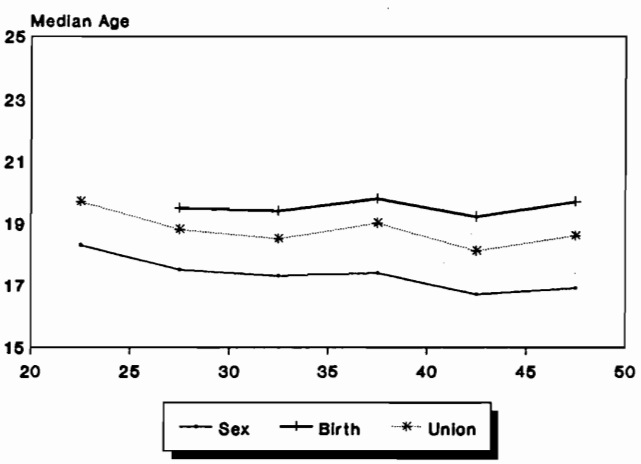
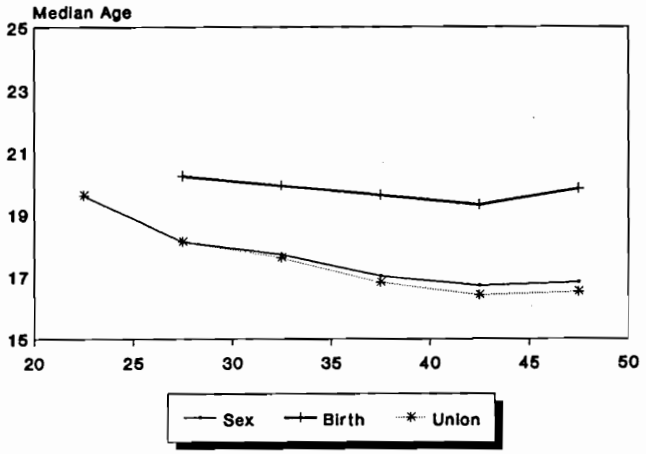


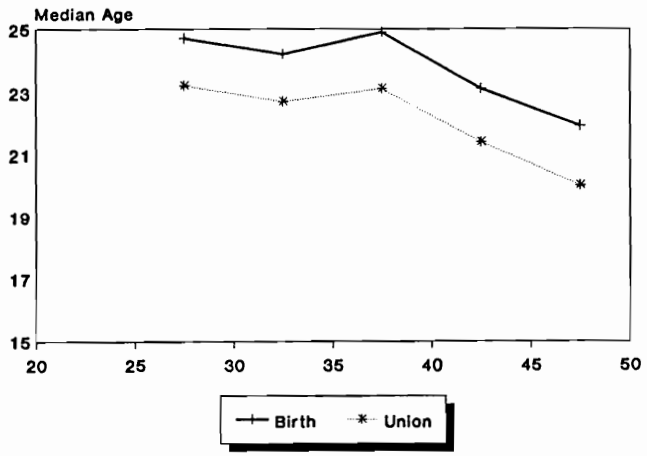
Figure 5.1—*Continued*

ASIA

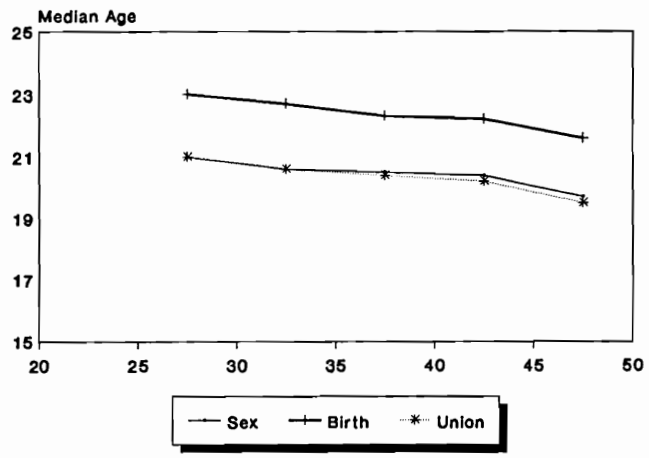
Indonesia



Sri Lanka



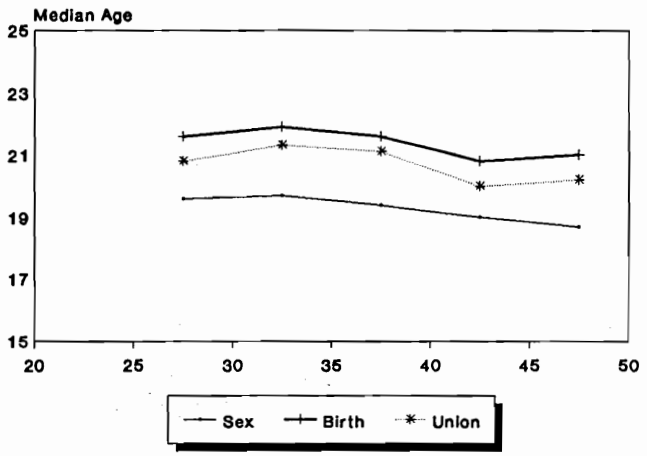
Thailand



LATIN AMERICA/CARIBBEAN

Brazil

Colombia

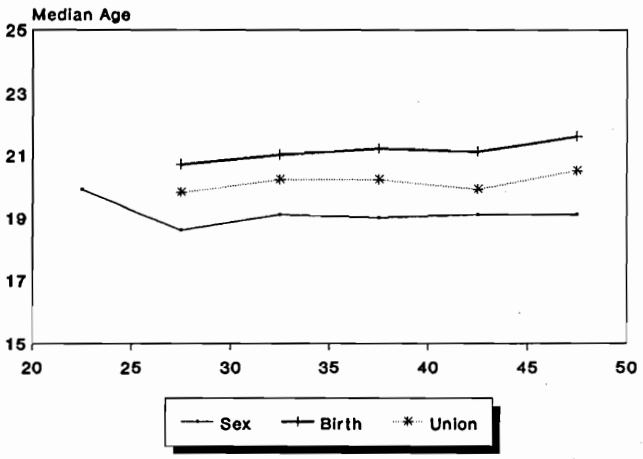


Dominican Republic

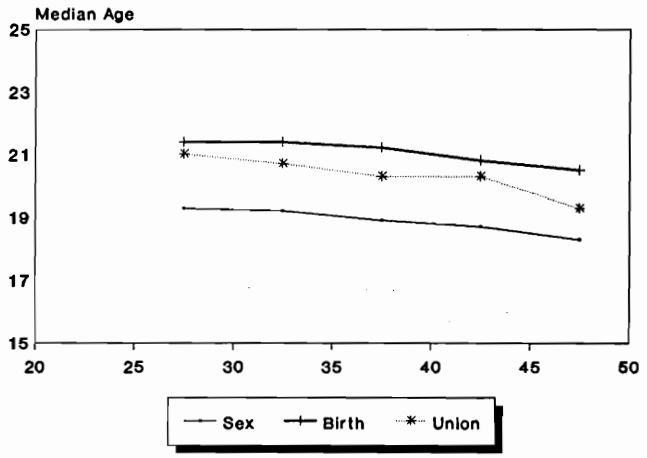
Figure 5.1—*Continued*

LATIN AMERICA/CARIBBEAN

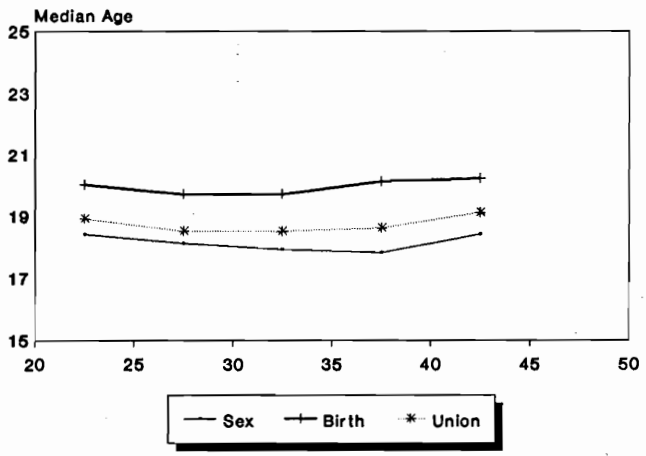
Ecuador



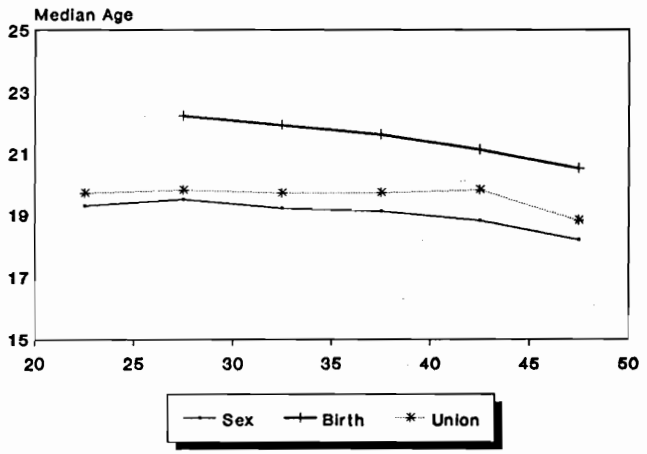
Peru



Guatemala



Trinidad and Tobago



Mexico

birth, for successive cohorts of women within each DHS dataset. This comparison shows how irregularities in one indicator are echoed in the others. For most of the countries, the indicators all increase or decrease about the same amount or stay constant. Some countries, notably Mexico and Togo, show a similar erratic pattern across age groups for age at first sexual intercourse, union, and birth. However, in Liberia, Uganda, and Trinidad and Tobago, the indicators do not share the same pattern. Age at first union and age at first birth show opposite trends in Liberia. In Uganda, age at first union and age at first birth, but not age at first intercourse, appear to have increased, lengthening the interval between the initiation of sexual intercourse and entry into a union. Possible underlying causes of this pattern, and other problems with the Liberia data, are discussed in more detail in the following paragraphs. In Trinidad and Tobago, the age at first sexual intercourse and union have remained constant but the age at first birth has increased steadily. Uganda and Trinidad and Tobago may be examples of countries in which women are postponing unions and/or births without changing their behavior in regards to initiating sexual relationships. These cases also demonstrate the value of using age at first sexual intercourse, rather than age at first union, as an indicator of the beginning of exposure to the risk of pregnancy.

The consistency of the data on age at first intercourse, age at first union, and age at first birth at the individual level is evaluated by examining the length of the intervals between these events. The percentage of negative first birth intervals ranges from zero in Morocco to 51 percent in Botswana. The percentages are highest in sub-Saharan Africa, lowest in North Africa and Asia, and moderate in Latin America and the Caribbean. These percentages are generally consistent with the proportion of women who report having sexual intercourse a year or more before their first marriage. Relatively large percentages of women report having sex for the first time a year or more after the start of their first union but the majority of these responses fall within one year of being consistent.

There are a few surveys with specific deficiencies which merit mention. The completeness of reporting of first unions and first births is significantly lower in Mali than in any other country. Only 6 percent of women reported a year and month of marriage and only 33 percent reported a year and month for their first birth. The substantial amount of imputation which was necessary in Mali suggests that caution should be used in relying on these data.

Information about age at first marriage and age at first birth for Liberia show a significant amount of heaping both on calendar years and durations. There is also some evidence that first marriages were displaced forward in the oldest cohort—or that there is some omission of first births among older women—and that first births were displaced forward by the oldest cohorts and backward by the youngest cohorts. Thus, trends over time in age at first marriage and age at first birth are likely to be distorted.

The data on first marriage in Botswana should also be mentioned. The percentage of first births which occurred prior to first marriage is very high in Botswana and it is the only country in which the median age at first marriage exceeds the median age at first birth. This is not necessarily an indication of deficiencies in the data, however. Stable sexual relationships that do not involve cohabitation are common in Botswana and many women have children in such unions. The data indicate that almost half of the women who reported themselves as never or formerly married were sexually active in the month before the interview (Lesetedi et al., 1989). Thus, in Botswana, the information on age at first intercourse is a better indicator of the beginning of exposure to childbearing than age at marriage.

Apart from the specific problems mentioned above, the evaluation suggests overall that the quality of the data on first union and first birth for women 20 to 44 is satisfactory in most of the surveys. A complete analysis of trends across age cohorts is difficult, however, due to misreporting in the oldest and youngest cohorts.

An important finding that emerged from this evaluation is that it is feasible to collect information on age at first sexual intercourse. The level of non-response to this question is very low and where responses are inconsistent with age at first union it seems likely to be the result of interviewer miscalculation rather than respondent misreporting. These data can be improved by reducing the extent to which interviewers calculate the age at first intercourse based on the date of first marriage. The level of inconsistency between age at first marriage and age at first intercourse could be reduced by the addition of a code—indicating that the respondent had sexual intercourse for the first time when she got married—to the question on age at first intercourse.

The information on age at first sexual intercourse also shows that in many countries there is a considerable amount of exposure to childbearing before the first union. In addition, trends in the age at which women are first exposed to the risk of pregnancy do not always parallel trends in the age at which women first enter a union. This information should be useful in analyses of age-specific fertility and marital fertility.

To improve the quality of data on date of first union it is recommended that the data collection procedures be modified slightly. Specifically, it is recommended that the data

be collected parallel to information on the respondent's age and birth date: each ever-married respondent is first asked her age (in completed years) at the time she began living with her first husband or partner; then, she is asked to provide the year and month she began living with him. As in the case of the respondent's age and birth date, the interviewer should be instructed to compare and correct the two pieces of information if they are inconsistent. Obtaining both date of and age at first marriage should also improve the quality of the data on age at first intercourse, since it will be easier for interviewers to calculate the correct age from responses such as, "I first had sex a year before I got married." The complexity of the calculation required of the interviewer in this case is obviously greater than in the case of the respondent's age and birth date. Careful training of interviewers will be necessary and some inconsistencies will undoubtedly remain in the questionnaires. Field editors and supervisors should be responsible for resolving these inconsistencies.

The magnitude of the data deficiencies highlighted in this report vary between regions and countries. Many of the surveys in Latin America, Asia, and North Africa exhibit no serious problems. Some of the surveys conducted in sub-Saharan Africa also appear to be relatively free of significant omission of births and misreporting of dates. Furthermore, improvements in the reporting of date of first marriage and age at first sexual intercourse can be achieved by implementing the procedures outlined above.

References

Coale, Ansley J. 1988. A Reassessment of Fertility Trends in Egypt, Taking Account of the Egyptian Fertility Survey. In *Demographic Responses to Modernization*, ed. Awad M. Hallouda et al., 21-86. Cairo, Egypt: Central Agency for Public Mobilisation and Statistics.

Goldman, Noreen. 1985. Assessment of the Fertility Data Collected in WFS Individual Surveys. In *Assessment of the Quality of Data in 41 WFS Surveys: A Comparative Approach*, Noreen Goldman et al., 38-62. WFS Comparative Studies, No. 44. Voorburg, Netherlands: International Statistical Institute.

Institute for Resource Development (IRD). 1989. *Demographic and Health Surveys Data Processing Manual*. Columbia, Maryland: IRD.

Lesetedi, Lesetedinyana T., Gaboratanelwe D. Mompati, Pilate Khulumani, Gwen N. Lesetedi, and Naomi Rutenberg. 1989. *Botswana Family Health Survey II 1988*. Columbia, Maryland: Central Statistics Office [Botswana] and Institute for Resource Development/Macro Systems, Inc.

Liskin, Laurie, N. Kak, A.H. Rutledge, L.C. Smit, and L. Stewart. 1985. Youth in the 1980s: Social and Health Concerns. *Population Reports*, Series M, Number 9. Baltimore, Maryland: Johns Hopkins University, Population Information Program.

Makinson, Carolyn. 1984. Age Overstatement Among Young Women and Its Effect on Estimates of Fertility and Proportions Married at Young Ages. Office of Population Research, Princeton University, Princeton, New Jersey.

Singh, Susheela. 1985. Assessment of Nuptiality Data. In *Assessment of the Quality of Data in 41 WFS Surveys: A Comparative Approach*, Noreen Goldman et al., 21-37. WFS Comparative Studies, No. 44. Voorburg, Netherlands: International Statistical Institute.

United Nations. 1987a. *A Comparative Evaluation of Data Quality in Thirty-Eight World Fertility Surveys*. New York: Department of International Economic and Social Affairs.

United Nations. 1987b. *Fertility Behaviour in the Context of Development: Evidence from the World Fertility Survey*. Population Studies, No. 100. New York: Department of International Economic and Social Affairs.

United Nations. 1988. *World Population Trends and Policies*. Population Studies, No. 103. New York: Department of International Economic and Social Affairs.

Assessment of the Quality of Birth History Data in the Demographic and Health Surveys

FRED ARNOLD

Acknowledgments

The author wishes to thank Guillermo Rojas for programming assistance. Thomas W. Pullum, Anne Cross, Noreen Goldman, Trevor Croft and Jeremiah Sullivan provided valuable comments and suggestions.

1 Introduction

Demographic estimates have improved immeasurably over the last several decades with improvements in survey design and demographic techniques. The latest innovations in survey research techniques and in demographic measurement have been incorporated in the Demographic and Health Surveys (DHS). Nevertheless, all survey data contain both sampling errors and non-sampling errors that can affect demographic estimates. This report focuses on non-sampling errors in DHS birth history data and their impact on current and past fertility estimates.

Reliable fertility estimates require complete and accurate reporting of the total number of women of childbearing age and the total number of children they have had. Information on the date of birth of the women and their children must also be accurately recorded. Fertility estimates may be affected by underreporting (or, less likely, overreporting) of the number of women or children, as well as a variety of errors regarding date of birth or age. This assessment of the quality of birth history data focuses on data that are relevant to the estimation of fertility rates. The report by Rutstein and Bicego in this volume evaluates errors in age reporting for women of childbearing age, and the reports on mortality (Sullivan et al.) and age at first birth (Blanc and Rutenberg) herein also contain information relevant to the assessment of birth reporting in DHS surveys.

The objectives of this evaluation are to examine a number of possible errors in DHS birth history data, assess the impact of these errors on fertility rates, and suggest ways of reducing or eliminating these problems in the future. The recommendations proposed entail modification of the DHS questionnaires, interviewer training, field procedures, and standard tabulations.

The first section examines the completeness of information on children's dates of birth and ages and the use of imputation when such information is incomplete or inconsistent. The next sections explore several types of age misreporting including age heaping and miscalculation or displacement of the date of birth. This is followed by a discussion of the possible underreporting of births. The final section compares DHS fertility estimates with estimates from other sources for overlapping periods. The overall conclusion is that the quality of the birth history data is generally good but that minor changes in survey methods could result in improvements in the estimation of fertility.

2 Completeness of Reporting of Children's Birth Dates

One way of evaluating the accuracy of information on children's birth dates is to examine how completely these dates were reported in the DHS birth histories and how many children were missing information that had to be imputed. Respondents were asked to give the year and month of birth of each of their children and, in addition, to supply the current age of each living child. When complete information on year and month of birth was not reported, the dates were assigned by imputation.[1] First, a logical range of possible values was constructed for each of these births and then a value was randomly assigned to produce a uniform distribution within that range. Partial information about the date of birth (such as the year of birth or the child's current age) was used to set strict constraints on the logical range of birth dates.

Additional constraints were imposed by assuming a minimum time gap between various events or activities. For example, births had to be at least seven months apart and a woman must have reached at least a minimum age (usually 12 years old) before having her first birth. Moreover, it is assumed that no child can be born less than seven months after a woman's age at first sex or at any time after the date of conception of a current pregnancy or after the date of sterilization, if any. Another set of constraints is imposed if responses are recorded for sections of the questionnaire that are limited to children of particular ages, such as health questions or height and weight measurements. Finally, ancillary information may be considered in the imputation process, such as the duration of contraceptive use, breastfeeding, postpartum amenorrhea, or abstinence. The imputation program has evolved during the course of the DHS surveys, so the procedures may differ slightly in different countries. The basic design, however, follows the process outlined above.

For each birth in the data file there is a date flag indicating whether or not the case was imputed and the type of information on which the imputation was based. Table 2.1 summarizes data on the completeness of information about the date of birth and age of each child. The interviewers or supervisors may have "imputed" some date of birth information in the field, when it was not provided by the respondent, but the extent of that practice is unknown. Table 2.1, therefore, includes only information on imputation carried out by computer.

[1] For a complete description of imputation procedures, see the DHS data processing manual (Institute for Resource Development, 1989).

Information on date of birth was recorded most completely in Latin America and the Caribbean, where 94-99 percent of all births did not require imputation. Even when the birth information was incomplete, it was usually only the month of birth that was missing. Virtually all of the cases in that region included information on either the year of birth or the current age of the child (or both). In Asia, the completeness of birth information varies greatly. Reporting deficiencies were substantial in Indonesia where 23 percent of births required some imputation. In Sri Lanka and Thailand, on the other hand, imputation was required for less than 10 percent of all births.

The situation in Africa also varies greatly. Botswana, Tunisia, Uganda, and Zimbabwe had levels of completeness of information comparable to Latin America, whereas Burundi and Morocco, and all the countries in West Africa, had substantial deficiencies in reporting of birth dates. The worst case was Mali where nearly two-thirds of the births required some imputation and 9 percent of the births had no information at all, either for the date of birth or the child's age. Even in Mali, however, more than 90 percent of children had at least a year of birth or age recorded. Therefore, if the partial information provided was reliable, the imputed dates of birth of these children should be accurate to within 12 months of their true dates of birth.

Overall, the completeness of reporting of children's birth dates is much better in DHS surveys than in comparable surveys carried out by the World Fertility Survey (WFS) in the same countries (see Table 2.1). The only exceptions are Morocco, where the completeness of reporting was similar in the two surveys, and Senegal, where the exceptionally high rate of completeness in WFS is an artifact of special data collection procedures. In Senegal, an age-event chart was used to record all births and interviewers were forced to impute the dates of an unknown number of births in the field (Chidambaram and Sathar, 1984).

The degree of improvement between WFS and DHS was dramatic in several countries in which the percent of children with imputed birth dates dropped by 15 or more percentage points. There are many possible reasons for this improvement. First, as levels of educational attainment improve, women are more likely to know the birth dates of their children. Second, in many societies, better knowledge of ages may be related to modernization and an increased emphasis on ages and timing of events in general. Finally, pressure was put on DHS interviewers to obtain a date of

Table 2.1 Percent distribution of children born to survey respondents by completeness of information on date of birth and age of children, Demographic and Health Surveys, 1986-1989

Country	No Imputation		Year and Age Reported, Month Imputed	Year Reported Month Imputed	Age Reported, Year and Month Imputed	Other Partial Information	All Imputed	Total
	DHS	WFS						
SUB-SAHARAN AFRICA								
Botswana	96.3		2.6	0.3	0.1	--	0.7	100.0
Burundi	78.6		13.7	7.4	a	a	0.2	100.0
Ghana	75.3	63.5	2.8	6.8	0.1	14.8	0.2	100.0
Liberia	83.7		8.2	5.4	1.1	1.2	0.5	100.0
Mali	34.9		--	14.0	5.7	36.3	9.2	100.0
Senegal	76.5	99.0	14.3	8.6	--	--	0.6	100.0
Togo	50.0		5.2	12.6	31.1	a	1.1	100.0
Uganda	99.9		a	a	a	a	a	100.0
Zimbabwe	99.4		0.1	0.2	0.1	--	0.2	100.0
NORTH AFRICA								
Morocco	57.1	59.7	2.2	1.0	0.7	38.6	0.5	100.0
Tunisia	94.8	70.4	1.3	1.8	0.7	--	1.4	100.0
ASIA								
Indonesia	77.2	46.5	16.2	6.6	a	a	0.1	100.0
Sri Lanka	92.2	73.4	4.0	0.9	2.0	--	0.8	100.0
Thailand	90.9	84.2	2.4	1.8	1.9	--	3.1	100.0
LATIN AMERICA/CARIBBEAN								
Brazil	96.7		--	2.2	--	0.2	0.9	100.0
Colombia	97.9	90.9	1.1	1.0	--	--	--	100.0
Dominican Republic	98.0	91.0	1.1	1.0	--	--	--	100.0
Ecuador	94.2	78.5	4.7	--	1.0	--	0.2	100.0
Guatemala	96.2		1.3	2.3	0.1	0.1	0.2	100.0
Mexico	98.4	NA	1.4	--	0.2	--	a	100.0
Peru	97.9	93.1	0.4	1.5	0.1	--	--	100.0
Trinidad & Tobago	99.0	94.5	0.2	0.5	--	a	0.3	100.0

Note: All figures are for DHS surveys except column 2.
NA = not available
[a] Less than 0.05 percent
Source: Chidambaram and Sathar (1984:31), Sepulveda (1984:13), Goldman et al. (1985:40), United Nations (1987:25).

birth for each child even if the respondent could not easily supply that information at first. Interviewers were required to check birth certificates or immunization records, if possible, or to try to determine a child's date of birth by comparison with firm dates for other children in the household. The interviewer's manual tells the interviewer that "You must enter a year [of birth], even if it is just your best estimate" (Institute for Resource Development, 1987:42). Because many DHS questions are asked only about children born after a specified date, special emphasis was placed on determining the date of birth in the field. All of these factors facilitated more complete reporting of birth dates in DHS than in WFS; however, it is not possible to compare the accuracy of the reported birth dates in the two sets of surveys.

Although the completeness of information on birth dates varies considerably from one country to another, the pattern of differentials according to the characteristics of respondents and their children is strikingly similar across countries (see Tables 2.2-2.4). Reporting of birth dates is less complete among older women and for births that occurred further in the past. In addition to the difficulty women may have remembering details about births that occurred many years ago, the reporting of those births is probably less complete because of the relatively low level of education of older respondents and because a larger proportion of early births did not survive. The relationship between completeness of reporting and education and child survival is shown in Tables 2.3 and 2.4, respectively. Completeness of reporting is positively related to education in every country. The reporting of dates of birth of children is complete or nearly complete among women with more than a secondary school education in every country except Burundi and Morocco. There is also a notable difference in all countries in the completeness of reporting

Table 2.2 Percentage of children with complete information on year and month of birth by number of years since birth of child, Demographic and Health Surveys, 1986-1989

Country	Number of Years Since Birth of Child								
	0	1	2	3	4	5-9	10-14	15-19	20+
SUB-SAHARAN AFRICA									
Botswana	99.9	99.6	99.2	98.9	98.9	97.2	94.4	93.6	92.2
Burundi	99.8	97.3	95.0	92.4	85.1	77.3	69.7	67.1	60.1
Ghana	96.6	92.0	87.1	85.6	81.2	74.4	72.0	66.1	62.1
Liberia	94.6	89.4	91.4	85.4	86.6	84.2	81.8	77.4	74.9
Mali	75.1	58.3	45.3	44.9	40.2	31.0	27.3	27.2	22.0
Senegal	99.0	97.0	93.5	91.5	90.9	78.8	66.0	61.3	57.6
Togo	86.1	82.6	76.1	67.4	58.9	49.9	40.8	33.8	22.3
Uganda	99.9	100.0	99.9	100.0	100.0	99.9	99.9	99.6	100.0
Zimbabwe	100.0	100.0	99.6	99.8	99.7	99.4	99.5	99.2	98.6
NORTH AFRICA									
Morocco	91.8	83.2	72.0	72.5	67.3	60.7	53.3	46.0	35.9
Tunisia	98.6	98.0	97.3	98.3	97.4	95.4	94.7	93.5	88.5
ASIA									
Indonesia	98.7	94.1	91.4	88.2	88.3	82.0	76.4	71.6	61.0
Sri Lanka	100.0	99.7	99.3	98.7	98.4	93.5	92.0	91.3	81.9
Thailand	99.5	98.1	96.9	95.8	96.5	93.8	90.7	88.9	82.1
LATIN AMERICA/CARIBBEAN									
Brazil	99.5	99.0	99.5	98.7	99.1	96.5	96.0	95.6	91.8
Colombia	100.0	99.6	98.7	99.3	99.7	98.6	97.4	97.0	96.6
Dominican Republic	100.0	99.9	99.7	99.6	99.3	98.2	97.6	96.9	96.8
Ecuador	98.9	98.4	98.8	97.2	96.9	94.9	93.7	91.6	88.3
Guatemala	99.9	99.5	98.8	98.5	97.4	96.5	94.9	93.3	91.4
Mexico	99.9	99.8	99.5	99.1	98.9	98.9	98.7	98.0	96.0
Peru	100.0	99.8	99.2	99.2	98.8	98.2	97.5	97.8	96.1
Trinidad & Tobago	100.0	100.0	100.0	100.0	99.5	99.5	99.3	98.9	97.1

of living and dead children. A simple (unweighted) average of the completeness of reporting in DHS countries shows that the year and month of birth are reported for 88.2 percent of living children and only 70.0 percent of dead children. Reporting is also more complete in urban areas than in rural areas in every country except Uganda, where complete reporting is nearly universal. Interestingly, there is virtually no difference in the completeness of reporting for male and female children. Where a difference exists, female children usually have a slight edge, perhaps because they are more likely to survive than males.

Table 2.3 Percentage of children with complete information on year and month of birth by age and education of mother, Demographic and Health Surveys, 1986-1989

Country	Age of Mother			Education of Mother			
	15-24	25-34	35+	None	Primary	Secondary	Higher
SUB-SAHARAN AFRICA							
Botswana	99.1	97.5	94.5	90.5	99.3	99.9	100.0
Burundi	96.8	85.1	71.8	77.3	83.2	97.5	76.2
Ghana	89.4	80.3	69.9	67.2	83.9	97.5	100.0
Liberia	91.5	86.3	79.3	80.2	90.9	96.3	99.2
Mali	49.8	39.5	28.6	32.4	54.9	92.8	(100.0)
Senegal	90.1	81.3	69.3	73.9	92.1	98.3	100.0
Togo	77.0	59.1	40.6	41.2	71.8	96.3	(100.0)
Uganda	99.9	100.0	99.8	99.9	99.9	100.0	(100.0)
Zimbabwe	99.8	99.6	99.2	98.3	99.7	99.9	100.0
NORTH AFRICA							
Morocco	69.0	64.0	52.2	52.8	84.6	97.1	93.9
Tunisia	96.9	96.7	93.6	93.1	98.5	99.6	99.2
ASIA							
Indonesia	93.0	85.5	70.1	60.8	80.5	95.6	98.6
Sri Lanka	99.1	95.9	89.7	78.8	89.7	97.6	99.3
Thailand	96.2	95.3	88.0	79.5	92.4	97.7	97.9
LATIN AMERICA/CARIBBEAN							
Brazil	99.3	97.2	95.8	92.1	97.3	100.0	100.0
Colombia	99.1	98.7	97.3	95.5	97.9	99.5	100.0
Dominican Republic	99.9	98.9	97.0	94.1	98.1	100.0	100.0
Ecuador	98.4	96.8	91.7	82.9	95.1	98.9	99.8
Guatemala	99.2	96.8	94.8	94.0	98.6	99.4	100.0
Mexico	99.2	99.2	97.7	95.6	99.1	100.0	100.0
Peru	99.4	98.9	97.3	93.9	98.8	99.9	100.0
Trinidad & Tobago	99.9	99.7	98.4	95.5	98.8	99.7	100.0

Note: Figures in parentheses are based on fewer than 25 cases.

Table 2.4 Percentage of children with complete information on year and month of birth by selected background characteristics, Demographic and Health Surveys, 1986-1989

Country	Survival Status of Child		Residence		Sex of Child	
	Dead	Living	Urban	Rural	Male	Female
SUB-SAHARAN AFRICA						
Botswana	90.4	96.7	98.6	95.5	96.1	96.4
Burundi	61.4	82.8	87.8	78.2	78.5	78.6
Ghana	59.3	78.2	82.9	72.2	75.3	75.4
Liberia	72.3	87.5	88.7	80.4	83.4	83.9
Mali	26.2	38.8	47.9	30.9	35.0	34.7
Senegal	61.6	81.2	88.2	70.3	76.2	76.8
Togo	26.5	55.3	67.5	44.2	50.7	49.3
Uganda	99.7	99.9	99.8	99.9	99.9	99.9
Zimbabwe	96.7	99.7	99.9	99.2	99.4	99.5
NORTH AFRICA						
Morocco	28.5	62.0	77.7	44.5	57.2	57.0
Tunisia	70.4	97.8	96.9	92.5	94.6	95.1
ASIA						
Indonesia	54.3	81.1	86.3	73.8	76.8	77.6
Sri Lanka	70.8	93.5	96.2	91.5	92.1	92.4
Thailand	43.6	95.0	91.1	90.9	90.5	91.4
LATIN AMERICA/CARIBBEAN						
Brazil	78.0	99.0	97.0	96.1	96.1	97.3
Colombia	87.7	98.8	98.3	97.4	97.9	98.0
Dominican Republic	91.6	98.8	98.7	97.1	97.8	98.2
Ecuador	73.3	97.0	97.6	90.8	93.9	94.5
Guatemala	81.5	98.5	97.8	95.5	96.2	96.2
Mexico	89.9	99.3	99.2	98.2	98.3	98.5
Peru	89.2	99.3	99.3	96.2	98.1	97.8
Trinidad & Tobago	86.9	99.6	99.1	99.0	99.0	99.0

3 Displacement of Children's Birth Dates

The DHS questionnaire contains many questions on fertility planning, health, and breastfeeding that are asked only about children who were born after a fixed cut-off date (usually January of the fifth year before the year of interview). Interviewers who want to decrease their workload may be inclined to change the birth dates of selected children so that the children will not be included in those sections of the questionnaire. This is particularly likely to occur if the respondent is unable to give an exact year of birth for all her children. Even when the date of birth is reported, however, interviewers may change the year of birth to avoid asking a large number of questions. This displacement problem is likely to increase in the second phase of the DHS program (DHS-II), since the sections of the core questionnaire that depend on children's year of birth are longer than in the previous core questionnaire. When displacement occurs, various estimates of fertility and mortality may be biased. It is difficult to avoid the problem entirely, but some measures to reduce its effects are outlined below.

3.1 DESCRIPTION OF THE PROBLEM

In the 22 DHS countries included in this report, an average of 43 questions about the last child were included in the sections with a filter for the date of birth of the child. In the same sections, an average of 29 questions were included for each of the previous children. The total number of such questions for previous children ranged from 13 in Liberia to 49 in Sri Lanka. Of course, every one of these questions is not asked about each eligible child, but in most cases, at least half of the total number of such questions would have to be asked. Moreover, interviewers are required to fill in, for each eligible child, a substantial amount of information from earlier questions. Therefore, interviewers can save time and effort by not including a child in these sections.

The problem is likely to be compounded in DHS-II because the new core questionnaire has additional questions that depend on a child's date of birth (see Table 3.1). For example, for the last birth since the cut-off date, the DHS-II B-core questionnaire (designed for use in low contraceptive prevalence countries) had more than twice as many questions as the DHS-I B-core questionnaire, plus about 10 more additional pieces of information that needed to be filled in by the interviewer. For previous births (which are more likely to be transferred out of these sections), the pretest questionnaire also had twice as many questions plus 6 more additional pieces of information than the DHS-I B-core questionnaire. Also, the extra burden of placing some of the information on the specially-designed DHS-II calendar (in countries in which the calendar is used) may further encourage interviewers to exclude children from the health and breastfeeding sections of the questionnaire. In the pretest in Trinidad and Tobago, the section of the questionnaire on health and breastfeeding, which is restricted to children born between January 1984 and July 1989, required a median of 13 minutes of interview time for women with one child born in that period, 21 minutes if there were two eligible children, and 26 minutes if there were three eligible children (Blanc et al., 1989). Therefore, interviewers could save a considerable amount of time and effort by moving children out of the eligible period. After the pretest, seven additional questions were added to that section of the core questionnaire, so the incentive for displacing births will be even greater than before. It should be noted that the displacement of birth dates in DHS-I was

Table 3.1 Number of questions in the DHS-I and DHS-II core questionnaires that depend on children's year of birth, Demographic and Health Surveys

Questionnaire	Questions About Last Birth	Questions About Previous Births
DHS-I model "B" questionnaire		
Questions asked of respondents	22	20
Other information filled in by the interviewer	17	17
Total	39	37
DHS-I model "B" questionnaire with additional health questions		
Questions asked of respondents	32	29
Other information filled in by the interviewer	18	17
Total	50	46
DHS-II model "B" questionnaire		
Questions asked of respondents	72	55
Other information filled in by the interviewer	28	23
Total	100	78

not strongly related to the exact number of questions asked about recent births, but the relationship was sufficiently strong to indicate the likelihood of greater problems in DHS-II.

3.2 EVIDENCE FOR THE DISPLACEMENT OF BIRTH DATES

It is difficult to measure the extent of displacement precisely, but examination of the year of birth distributions of children helps to identify countries in which displacement is a significant problem. Table 3.2 shows the annual number of births recorded in the year of the survey and in the nine years before the survey. Children born in the fifth year prior to the survey are the oldest children included in the health, breastfeeding, and fertility planning sections of the questionnaire. That is, the column "five years prior to the survey" contains all of the children born in the earliest 12 months included in those sections. If births are being incorrectly transferred from that year to the previous year, then a shortage of births should be evident in that column and an excess of births should appear in the next column.

Table 3.2 Number of births by calendar years prior to the survey, Demographic and Health Surveys, 1986-1989

	Years Prior to the Survey[a]									
Country	0	1	2	3	4	5	6	7	8	9
SUB-SAHARAN AFRICA										
Botswana	593	681	594	630	600	627	742	559	641	567
Burundi	361	869	805	762	734	669	880	736	635	597
Ghana	241	888	850	843	787	791	782	688	713	707
Liberia	477	1296	942	1002	984	778	1216	916	843	741
Mali	182	841	625	623	660	597	763	804	698	670
Senegal	159	937	857	848	869	794	921	899	797	702
Togo	164	674	654	570	655	534	665	656	543	512
Uganda[b]	1006	1075	1027	982	899	881	932	736	861	743
Zimbabwe[b]	549	685	715	639	722	697	712	582	666	572
NORTH AFRICA										
Morocco	667	1250	1211	1223	1262	1190	1339	1353	1250	1141
Tunisia	221	932	856	893	942	861	945	879	874	808
ASIA										
Indonesia	1385	1442	1644	1842	1744	1991	1864	2028	1876	1750
Sri Lanka	106	788	799	794	839	810	892	921	838	848
Thailand	215	748	801	661	781	769	812	837	901	745
LATIN AMERICA/CARIBBEAN										
Brazil	363	639	663	753	737	730	763	611	681	574
Colombia	435	558	548	513	577	556	577	584	540	495
Dominican Republic	706	943	815	909	902	862	822	904	816	770
Ecuador	82	678	622	562	546	664	634	687	586	603
Guatemala	871	919	926	892	937	835	1041	889	840	734
Mexico	326	1237	1064	1191	1090	1145	1069	1131	1109	1048
Peru	529	597	621	637	686	677	679	613	667	599
Trinidad & Tobago	188	386	390	405	404	371	410	387	356	323

[a]Births are for calendar years for all countries with a January cut-off date for the fertility planning and health and breastfeeding sections of the questionnaire. For Mali, Senegal, Tunisia and Togo, which had different cut-off dates, births are for 12-month periods starting in March, April, May and June, respectively.

[b]For Uganda and Zimbabwe, zero years prior to the survey refers to 1988. The table excludes 8 children in Uganda and 2 children in Zimbabwe who were born in January 1989.

The first two columns of Table 3.3 show the relationship between births in the fifth and sixth years prior to the sur-survey and the average number of births in the preceding and succeeding years. The value of these "birth year ratios" (which are similar to age ratios) would be expected to be approximately 100 in the absence of birth year displacement, heaping on particular years of birth, or erratic annual changes in the total number of births. Displacement would be suspected if the value of the ratio in column one of Table 3.3 is substantially less than 100 and/or if the value of the ratio in column two is substantially more than 100.

There is little evidence of displacement in most of Asia, Latin America, and the Caribbean, except for Guatemala and (to a lesser extent) Trinidad and Tobago. On the other hand, displacement is present in the majority of African countries. The most severe problem is in Liberia where more than 200 births may have been transferred from 1981 to 1980. Botswana, Burundi, Mali, and Togo also appear to have a substantial displacement problem. A smaller degree of displacement is evident in Morocco, Senegal, Tunisia, Uganda, and Zimbabwe. In two countries with displacement (Liberia and Senegal), the sixth year prior to the survey is 1980, which may receive extra births because of heaping on year of birth. For 1970, however, no similar heaping was found for births in Senegal and no heaping was found in Liberia for living children, who account for most of the displacement. The shortage of births five years before the survey is shown graphically in Figure 3.1 for the three countries in which displacement was most severe (Burundi, Guatemala, and Liberia). For comparative purposes, Figure 3.1 also includes data for Colombia, a country in which no displacement is evident.

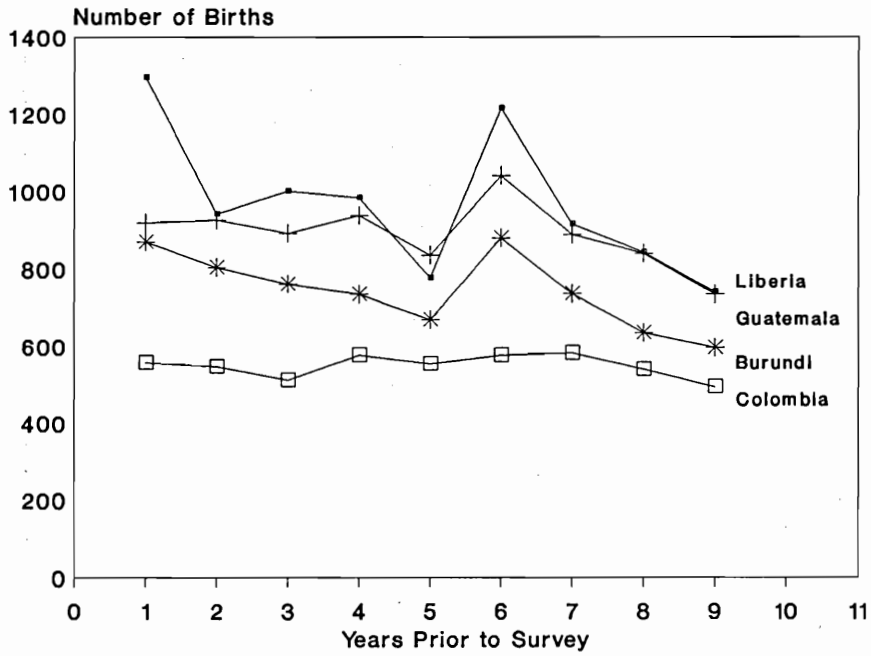
Table 3.3 Birth year ratios by survival status of births, Demographic and Health Surveys, 1986-1989

	Centered on Period:		Centered on Period 5 Years Before the Survey	
Country	5 Years Before the Survey	6 Years Before the Survey	Dead Children	Living Children
SUB-SAHARAN AFRICA				
Botswana	93.4	125.1	66.7	95.1
Burundi	82.9	125.3	85.7	82.5
Ghana	100.8	105.7	107.3	99.7
Liberia	70.7	143.6	71.2	70.6
Mali	83.9	108.9	62.5	91.2
Senegal	88.7	108.8	83.1	90.0
Togo	80.9	111.8	61.8	84.5
Uganda	96.2	115.3	84.9	98.9
Zimbabwe	97.2	111.3	126.0	95.1
NORTH AFRICA				
Morocco	91.5	105.3	95.7	91.0
Tunisia	91.3	108.6	81.4	92.0
ASIA				
Indonesia	110.4	92.8	113.8	110.0
Sri Lanka	93.6	103.1	75.7	94.4
Thailand	96.5	101.1	114.3	95.6
LATIN AMERICA/CARIBBEAN				
Brazil	97.3	113.8	73.2	99.9
Colombia	96.4	101.2	106.1	95.7
Dominican Republic	100.0	93.1	72.0	103.2
Ecuador	112.5	93.9	93.2	114.4
Guatemala	84.4	120.8	71.7	86.1
Mexico	106.1	93.9	101.4	106.2
Peru	99.2	105.3	80.5	101.7
Trinidad & Tobago	91.2	108.2	33.3	93.8

Note: The birth year ratio x years before the survey = $\frac{B_x}{.5(B_{x-1} + B_{x+1})} \times 100$,

where B_x = number of births x years before the survey.

Figure 3.1—Number of births by calendar year, selected Demographic and Health Surveys, 1986-1987



It might be expected that displacement would be more common for living children than for dead children since more questions are asked about living children. The opposite is true, however, in most countries (see the last two columns of Table 3.3). In the majority of countries with a displacement problem, the problem is greater for dead children than for those who are living. Moreover, displacement of dead children is evident in some countries that do not exhibit any displacement of living children (for example, the Dominican Republic and Peru). Three reasons may be suggested for the prevalence of displacement for children who have died. First, interviewers may be embarrassed to ask a series of detailed questions about children who are no longer living. Therefore, the incentive to transfer deceased children across the birth date cut-off may be high even though fewer questions would need to be asked about such children. Second, information on the date of birth is likely to be less precise or less certain for dead children so that interviewers have more leeway in estimating their dates of birth. Finally, when the year of birth is unknown for a child who died, the data entry and imputation programs generally assume that the child is too old to be included in the health and breastfeeding section of the questionnaire unless the respondent had a previous birth that is eligible for those sections. This problem is described more fully in the article on mortality in this volume.

3.3 IMPLICATIONS

The effect of displacement of birth dates on demographic estimates depends on both the magnitude of the displacement problem and the type of displacement that occurs. If interviewers are simply moving births backwards in time by a few months, the impact will be smaller than if they are moving them back a whole year. The data are inconclusive on this point. It should be noted, however, that almost all countries with substantial displacement have much less distortion in their single year of age distributions than in their single calendar year of birth distributions. This would seem to indicate that much of the transference across calendar years is taking place within a single year of age. Displacement of this type will have less impact on many demographic measures.

Demographic rates will not be affected unless a child is transferred across the same boundary that is used for reporting those rates. In the DHS surveys, fertility rates are most commonly reported for each five-year period before the interview date and for the two most recent three-year calendar periods before the survey (where the most recent period includes fertility in the year of the survey plus the three full years before the survey). Displacement out of the fifth calendar year before the survey will have no effect on fertility estimates for any of the

three-year periods. Significant displacement, however, will cause fertility to be underestimated in the most recent five-year period and overestimated in the preceding five-year period.

This effect is particularly noticeable in surveys with fieldwork in the early months of the year. The timing of the survey is important because children born in the earliest year of eligibility are either four or five years old at the time of the survey. The nature of displacement, then, depends on what proportion of children are age four and what proportion are age five before any displacement takes place. For surveys conducted early in the year, the majority of these children are four years old and their age is likely to be changed to five. For surveys conducted late in the year, the majority of these children are five years old and their age is likely to be changed to six. Therefore, displacement should have virtually no effect on the standard measures of fertility for surveys conducted at the end of the year, such as those in Guatemala and Botswana. The effect may be quite large, on the other hand, in a country such as Liberia where the fieldwork was conducted primarily from March through June. This undoubtedly explains why the most recent five-year total fertility rate for Liberia (6.3) is out of line with the rates for 1980-82 (6.7) and 1983-86 (6.5). In fact, it is estimated that the TFR for the most recent five-year period would have been 6.5 instead of 6.3 if there had not been any displacement of birth dates. This estimate was made by randomly reallocating 219 weighted births in 1980 to the year 1981 so that there would be an equal number of births in those two years.

It should be pointed out that, despite the magnitude of displacement in Liberia, its effect on the five-year TFR is quite modest. The impact is more evident when examining changes in fertility over time, however. Cumulative fertility at ages 15 to 39 in Liberia was originally estimated to have declined from 6.2 children in the period 5-9 years before the survey to 5.7 children in the period 0-4 years before the survey. After adjusting for displacement, however, the fertility decline virtually disappears (cumulative fertility remains constant at 5.9 children for both periods).

Birth intervals will also be biased when there is displacement in children's dates of birth. Specifically, the preceding birth interval will be underestimated for displaced children and the succeeding birth interval will be overestimated. In fact, one way of detecting displacement would be to examine the average birth intervals of children by single calendar years of birth. The analysis of birth intervals is particularly useful for detecting displacement, since it is valid even if there are real fluctuations in the annual number of births. The birth year ratios discussed above, on the other hand, are affected by such annual fluctuations, which could complicate the analysis of birth displacement.

Figure 3.2 shows the average interval between births in each calendar year and the next succeeding birth in Burundi, Guatemala, and Liberia. In general, going back in time, a pattern of rapidly increasing birth intervals would be expected in the years immediately prior to the survey (because the time available to have another birth is constrained by the date of the survey) followed by a flat or gradually increasing trend. All three countries, however, exhibit somewhat higher than expected values six years before the survey and somewhat lower than expected values seven and eight years before the survey. This pattern is consistent with the displacement of some births from five years before the survey to six years before the survey. If the births that are displaced are randomly selected from those born five years before the survey, there would be no effect on the average birth intervals five years before the survey, but the average birth intervals six years before the survey would be overestimated. Moreover, the succeeding birth interval would be underestimated for the birth prior to the one that was displaced. Since the average birth interval is about two and one-half years at that point, that would imply an underestimate of birth intervals for births that occurred 7 to 8 years before the survey.

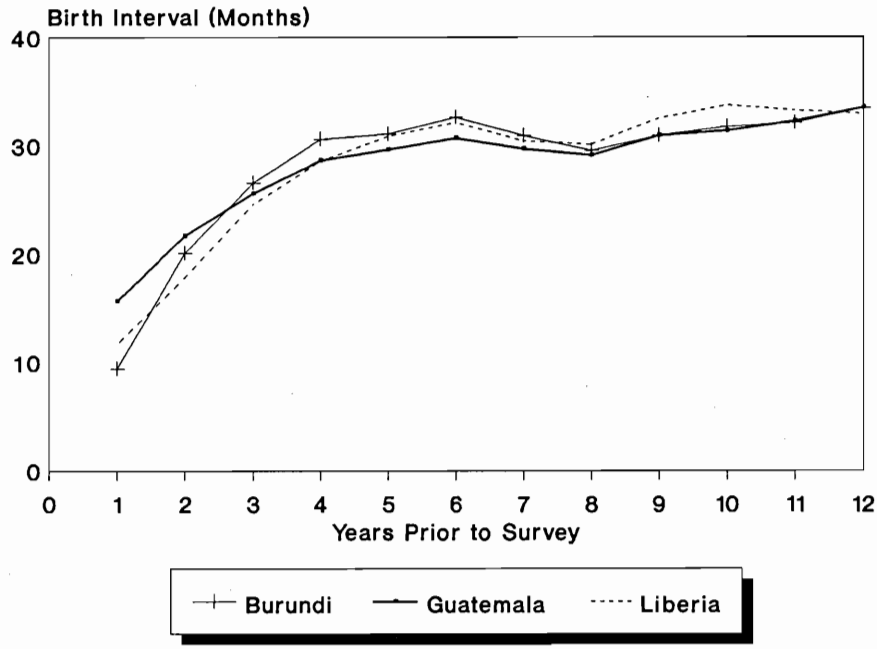
3.4 POSSIBLE SOLUTIONS TO THE PROBLEM

Various methods may be suggested to help minimize the problem of the displacement of birth dates and to reduce the impact when it cannot be eliminated. These methods include changes in interviewer training and field procedures, standard tabulations, and questionnaire design.

(1) Interviewers should be carefully trained to reduce displacement and various types of age errors. Even more crucial, however, is the close supervision of interviewers in the field. Preliminary analysis indicates that although the displacement of birth dates is common, some interviewers are particularly likely to commit abuses. Therefore, supervisors need to detect the problem early in the fieldwork, so that interviewers who are responsible for displacement of birth dates can be identified and supervised more closely.

Supervisors should pay particular attention to questionnaires reporting children born in the year or two before the cut-off date for the health and breastfeeding section. Whenever possible, it is desirable for supervisors to personally check the age information for all children with recorded dates of birth in that period. At a minimum, extensive spot-checking should be done for those births, particularly in the early part of the fieldwork. Displacement may also be detected through reinterviews, which are expected to be included in many DHS-II surveys. Another method of detection would be to tabulate the single-year-

Figure 3.2—Average succeeding birth interval by year of birth, selected Demographic and Health Surveys, 1986-1987



of-age distribution of children for each interviewer during the fieldwork. This could even be done in the field with laptop computers (if available). This type of tabulation, however, can only detect the most severe cases of displacement and, normally, the results would not be available early enough to be of use.

(2) Different cut-off points should be used for the eligibility of children and for reporting demographic rates. This objective can be accomplished by either changing the eligibility cut-off point for the health and breastfeeding section of the questionnaire or changing the periods of time over which demographic rates are calculated. Changing the questionnaire cut-off point from five years to six years would only shift the bias in demographic measurement from the five-year rates to the three-year rates. A cut-off at four years before the interview would have the least effect on both the three-year and five-year rates but it would reduce the sample of children with health information considerably. An alternative would be to reduce reliance on five-year demographic rates and focus on three-year rates and four-year rates. Although it is convenient to use five-year rates for the period-cohort fertility tables, there is no reason that mortality rates and total fertility rates cannot be reported for four-year periods instead of five-year periods when analyzing trends over time. Doing so would slightly increase the sampling error of the estimates, but that would be more than offset by the reduced bias in the estimates for many countries. Because of the displacement problem, four-year rates were chosen for analyzing fertility trends in the DHS comparative report on fertility (Arnold and Blanc, 1990).

(3) The number of questions in the health and breastfeeding section of the questionnaire should be reduced. As pointed out earlier, however, the relationship between displacement and the number of questions about eligible children was not strong in DHS-I. Because displacement was both widespread and serious in the case of dead children, consideration should be given to reducing the number of questions that are asked about children who have died. Since displacement is not a universal problem with respect to living children, however, reducing the number of questions for all children would be counterproductive in many countries.

4 Age Heaping

Countries in which age data are of poor quality often exhibit a preference for ages ending in certain digits and an avoidance of ages ending in other digits. Typically ages in ending zero and five will be overrepresented in age distributions for these populations and even-numbered digits may also be favored (Stockwell, 1966; Turner, 1958; Weller et al., 1987). Digit preference occurs most frequently among older individuals but it can also be found in the reported age distributions of children (Ewbank, 1981; Nagi et al., 1973). The nature and extent of age heaping can be evaluated by examining the age ratios centered on single years of age. Table 4.1 shows the age ratios for surviving children up to age 15, as reported in DHS birth histories. Small deviations of age ratios from 100 should not be interpreted as evidence of age heaping since they could be caused by real fluctuations in births or deaths. Larger deviations should be regarded as suspect unless they can be traced to specific historical events that could have influenced the annual number of births.

In general, there are few instances of serious age heaping for children in DHS surveys, although some distortions in the age distributions are evident. The most pronounced heaping is on age 10 with more than half of the countries having age ratios greater than 105 at that age. Heaping on age 10 is seen most frequently in sub-Saharan Africa, particularly in Liberia and Senegal. Several countries (including Burundi, Ecuador, Ghana, Guatemala, Peru, and Zimbabwe) exhibit a distinct preference for ages ending in even numbers. This result is consistent with the fact that censuses and surveys in Africa, Asia, and Latin America typically show a preference for even-numbered ages for children, particularly children over five (Ewbank, 1981; Scott and Sabagh, 1970).

Preference for ages ending in five is rare. There are at least three explanations for this situation. First, the preference for even digits in several countries makes heaping on five unlikely. Second, there may be some displacement from age 15 to age 14 for female children to avoid eligibility for the individual interview. Finally, the year of birth displacement discussed earlier will tend to shift some children from age five to age six, although there may also be a shift from age four to age five in countries in which the fieldwork was conducted early in the year. In fact, the lack of heaping on age five suggests that displacement of the birth year is a real phenomenon and not just a result of digit preference. In the annual series of births by calendar year, births are about equally likely to be underrepresented as overrepresented in years ending in five. Therefore, neither ages nor calendar years exhibit a proclivity for five as a final digit.

Digit preference of the magnitude exhibited for children in DHS surveys will not have a major impact on the estimation of demographic rates. Heaping on age 10 will tend to slightly underestimate fertility rates 5-9 years before the survey and slightly overestimate fertility rates 10-14 years before the survey. Thus, any decrease in fertility between these two periods will be similarly overestimated. A small effect will also be evident on the change in fertility rates between the first two 5-year periods before the survey date. Heaping on even-numbered ages will tend to overestimate fertility 0-4 years and 10-14 years before the survey and underestimate fertility 5-9 years before the survey (since some children who are actually age 5 or age 9 will be reported as age 4 or age 10, respectively). The effects will be quite modest, however. When 3-year calendar periods are used for reporting fertility rates (with the most recent period also including the year of the survey), the effect of a preference for even digits will generally be small and the precise type of bias will depend on the timing of the fieldwork during the calendar year.

In summary, while digit preference exists to some extent in most DHS surveys, it is not a major problem in the reporting of children's ages. Moreover, the impact of age heaping on fertility rates is quite small. Efforts have been made in all DHS surveys to obtain the exact calendar year and month of birth of children. Further improvements in the accuracy of age information may be possible, but some degree of age heaping is inevitable in countries in which mothers do not know their children's date of birth.

Table 4.1 Age ratios for living children by single year of age, Demographic and Health Surveys, 1986-1989

Country	Age Ratio Centered on Age[a]: 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
SUB-SAHARAN AFRICA															
Botswana	95.1	93.5	106.0	94.5	94.4	126.5	80.1	113.6	95.4	99.8	96.1	107.5	100.5	99.7	97.0
Burundi	85.3	117.9	90.7	102.1	93.4	111.7	102.6	92.6	94.0	121.9	74.5	119.5	91.5	114.5	87.3
Ghana	101.4	104.4	93.2	106.3	95.6	103.3	91.1	114.4	90.4	116.4	78.8	118.7	97.5	92.4	109.2
Liberia	85.9	89.5	111.6	91.5	103.8	104.9	97.7	101.5	87.8	131.1	78.1	106.7	93.1	110.4	99.4
Mali	88.1	96.8	100.6	102.7	90.3	116.7	91.0	107.8	90.2	116.1	87.9	101.1	101.1	110.9	84.9
Senegal	105.1	91.5	107.7	94.2	99.6	108.8	94.7	110.3	81.9	129.5	76.8	111.1	97.2	108.4	89.8
Togo	97.0	100.9	99.3	102.3	91.6	112.5	96.2	96.3	100.6	92.6	93.0	112.0	99.1	93.8	101.7
Uganda	101.4	96.3	104.7	95.0	94.2	123.8	74.0	121.0	94.1	109.8	82.6	117.3	91.9	108.5	88.5
Zimbabwe	99.6	109.2	89.0	108.7	98.2	105.7	89.7	116.4	90.3	101.4	93.2	105.9	104.5	97.2	102.0
NORTH AFRICA															
Morocco	98.5	94.2	108.2	98.6	89.6	113.9	100.5	94.0	104.0	99.9	96.3	104.0	99.5	100.8	96.8
Tunisia	95.7	103.4	98.1	106.8	90.7	108.7	95.2	107.9	93.6	101.7	100.3	99.3	103.9	92.9	113.8
ASIA															
Indonesia	92.0	100.2	108.4	92.5	104.4	99.5	105.5	97.5	95.1	108.0	90.3	111.5	96.2	99.5	104.0
Sri Lanka	103.3	97.3	102.0	98.7	96.4	107.1	99.8	100.6	97.4	104.6	95.2	104.1	98.0	98.3	102.1
Thailand	108.8	97.5	96.6	100.6	100.7	99.9	106.4	94.0	96.6	108.3	95.5	106.4	93.7	99.6	105.0
LATIN AMERICA/CARIBBEAN															
Brazil	92.3	92.9	118.1	91.2	104.6	103.4	93.2	106.7	94.3	105.2	95.2	105.0	91.9	114.1	91.9
Colombia	110.1	96.8	95.1	106.0	101.0	96.8	101.7	104.1	95.7	104.0	97.4	101.7	92.7	100.7	105.1
Dominican Republic	105.1	91.9	104.4	99.7	100.2	97.6	104.4	100.4	92.4	109.8	96.8	100.4	96.9	111.9	84.5
Ecuador	102.5	96.1	89.4	111.9	94.2	107.7	97.8	101.8	94.7	110.0	87.9	114.5	88.6	113.0	91.3
Guatemala	93.2	103.1	96.7	105.1	89.7	117.2	92.6	106.9	89.7	112.8	88.6	109.8	91.8	114.0	84.7
Mexico	94.5	102.8	96.7	104.4	99.3	97.0	102.9	101.9	94.5	106.1	95.7	103.1	99.2	101.4	101.0
Peru	92.9	104.1	94.8	105.7	100.0	104.8	89.1	115.0	85.4	120.3	83.7	117.1	91.7	106.8	91.5
Trinidad and Tobago	104.1	97.5	104.6	99.3	94.1	112.3	94.7	103.4	95.6	100.5	96.6	107.3	92.0	105.1	100.4

Note: Refers to children listed in birth histories who are still living.

[a] Age ratio at age x = $\frac{P_x}{1/2(P_{x-1} + P_{x+1})} \times 100$

5 Miscalculation of Year of Birth

In assessing the quality of DHS age data, most of the attention so far has been focused on age displacement and age heaping. Another potentially serious problem in the age data involves the incorrect calculation of the year of birth from the child's age by either the interviewer or the respondent. An examination of month of birth tabulations gives some indication of the extent of this problem.

5.1 DESCRIPTION OF THE PROBLEM

When the age of a child is known but the year of birth is unknown, the interviewer or the respondent may try to estimate the year of birth in the field by subtracting the child's age from the year of the interview. This calculation will result in the correct year of birth only if the month of interview is later than the month of birth. If the month of interview is earlier than the month of birth, then the calculated year of birth will be overestimated by one year. If the month of interview equals the month of birth, then this calculation will be correct only half of the time on average. A similar type of problem would occur if the year of birth is reported but the age is unknown. In that case, subtracting the year of birth from the year of interview will produce the correct age only if the child has already had a birthday in the year of interview.

In either case, the problem can be easily detected when the month of birth of a child is not reported. In that case, subtracting the child's age from the year of interview to obtain the year of birth may result in a very uneven distribution of imputed months of birth. Under such circumstances, the DHS imputation program forces the imputed month to fall in the part of the year before the individual interview was conducted. Although this procedure is logically correct and it produces consistent data, in many countries it causes imputed events to be bunched in the first part of the year. Moreover, the distortion is more serious the earlier in the year the fieldwork was conducted. This problem has become apparent in a number of DHS countries, particularly in Africa. For this reason, the imputation program was modified for Ghana and Mali (where the problem was particularly apparent) but no attempt was made to go back and recalculate the imputed values for the countries that had already been processed.

The effects of the imputation problem are most evident in Senegal. In that country, 96.8 percent of all cases in which the month of birth was imputed (and both the year of birth and age were recorded) had months of birth in the same month as the interview or earlier in the year (see Table 5.1). If births had been spread evenly throughout the year, on the other hand, only 44.2 percent of all births would have occurred in the month of interview or earlier. This indicates that, for many children, the interviewers or respondents were obtaining the year of birth by subtracting the child's age from the year of interview.

The distribution of imputed month of birth is also highly distorted for Burundi, Liberia, Togo, and Tunisia. This type of error was also found to be severe in the Ghana survey, although it is not apparent in Table 5.1 because the imputation program was modified to counteract the problem. Several other countries exhibit smaller amounts of distortion in the distribution of month of birth, although in some countries the number of cases that had the month of birth imputed is too small to attribute the distortion to erroneous calculation of the year of birth. Table 5.1 also indicates that there was relatively little distortion in countries where the fieldwork was conducted in the last few months of the year (Botswana, Colombia, the Dominican Republic, Guatemala, Indonesia and Peru). Curiously, the distortion is in the opposite direction in the only two countries in which the fieldwork was conducted in the first quarter of the year (Ecuador and Sri Lanka). One apparent reason for this anomaly is that ad hoc imputation programs for those countries were incorrectly assigning births to the months of January and December only half as often as they should have because of an error in the random number generator in ISSA (the Integrated System for Survey Analysis), the program used to process DHS data.

Some distortions also occur in the distribution of month of birth for children with no imputed birth information, but the degree of distortion is limited (see the last three columns of Table 5.1). The actual value is within five percentage points of the expected value in all countries except Burundi, Ghana, Liberia, and Tunisia. There is very little distortion in most of Asia and Latin America. The reason for the discrepancy in the African countries is unclear, but there are several possibilities. First, there is a tendency in some countries for births to be reported as occurring disproportionately in the same month as the interview was conducted. Whereas, on average, 8.3 percent of all births would be expected to occur in the same month as the month of interview, the actual values were 10.3 percent in Burundi, 10.2 percent in Ghana, 9.8 percent in

Table 5.1 Percentage of children whose month of birth falls in the month of interview or earlier, Demographic and Health Surveys, 1986-1989

Country	Imputed Cases[a]		Number of Children	Non-Imputed Cases		Number of Children
	Actual	Expected		Actual	Expected	
SUB-SAHARAN AFRICA						
Botswana	94.9	82.7	291	82.9	78.9	10851
Burundi	92.6	46.0	1649	51.5	43.6	9425
Ghana	27.6	30.1	2505	37.2	30.9	10706
Liberia	67.8	36.9	1336	45.5	39.1	13675
Mali	35.3	39.4	4600	36.6	36.2	4424
Senegal	96.8	44.2	2053	52.7	49.1	11008
Togo	90.3	69.5	557	71.2	69.3	5389
Uganda	b	b	6	79.1	78.7	16507
Zimbabwe	(94.4)	(85.6)	18	80.3	79.0	12333
NORTH AFRICA						
Morocco	66.3	52.1	552	56.3	51.3	14565
Tunisia	92.5	56.3	213	71.0	62.4	15612
ASIA						
Indonesia	99.6	87.3	6533	89.9	87.3	31171
Sri Lanka	4.5	14.6	707	18.1	17.1	16269
Thailand	47.9	37.1	444	38.7	36.0	16914
LATIN AMERICA/CARIBBEAN						
Brazil	55.2	53.9	260	55.7	55.1	11586
Colombia	98.6	88.7	126	87.4	87.4	11377
Dominican Rep.	99.6	90.0	196	81.2	81.7	17950
Ecuador	7.4	19.1	552	16.5	18.2	11144
Guatemala	99.5	90.2	183	91.0	90.2	14139
Mexico	39.8	32.7	321	32.7	33.3	23240
Peru	(100.0)	(92.1)	59	91.5	90.9	13016
Trinidad & Tobago	(83.3)	(64.4)	18	56.1	58.3	7760

Note: Figures in parentheses are based on fewer than 100 cases.
[a] Imputed births are those for which the month of birth was imputed but both year of birth and age were reported.
[b] Fewer than 10 cases

Liberia, 10.0 percent in Morocco, and 9.6 percent in Senegal. Second, interviewers who have been trained that the year of birth equals the year of interview minus age only when the month of birth is before the month of the interview may impute a month of birth early in the year to simplify their calculations. Finally, the distortion in the distribution of month of birth may be attributable in part to the seasonality of births in some African countries.

5.2 IMPLICATIONS

A one-year advance in the year of birth of some children will produce a bias in estimates of fertility rates and birth intervals. The birth intervals between a child with an imputed month of birth and the previous child may be overestimated, perhaps by as much as 12 months. Miscalculating the year of birth will not have much of an effect on estimates of fertility based on children one year of age and over, since approximately as many children will be erroneously moved into an age group as erroneously moved out of it. The effect will be greater on children who have not yet reached their first birthday, since children are only moved into the year of interview, not out of it. To the extent that there is a bias, fertility will be overestimated in the year before the survey (and, to a lesser extent, in the three years or five years before the survey). Similarly, any fertility decline over time will appear to be less steep than is actually the case. However, the effect on recent fertility is likely to be minimal since the year of birth is usually known for young children and there is relatively little imputation at early ages. For example, in Senegal 23.5 percent of all children have some imputation compared with 5.6 percent of children age 0-4 and 1.0 percent of children under age one. In Liberia, imputation was carried out for 16.3 percent of all children, 9.2 percent of children age 0-4, and 5.4 percent of children under age one. Similar patterns were found in other DHS countries (see Tables 2.1 and 2.2).

If a child's age is miscalculated from the given year of birth, on the other hand, the effects will be opposite those

noted above. This error would cause an underestimation of the preceding birth interval and an underestimation of recent fertility, as well. Although the data themselves do not indicate which type of miscalculation is more prevalent, field experience suggests that a respondent is more likely to report age only than year of birth only. Therefore, miscalculation of year of birth from reported age is probably the more serious problem.

5.3 POSSIBLE SOLUTIONS TO THE PROBLEM

Changes in field procedures, interviewer training, data processing procedures, and questionnaire design can be implemented to deal with the problem of incorrect calculation of the year of birth. Disadvantages associated with some of these changes will need to be taken into account, however. Some procedural modifications designed to alleviate the age calculation problem are described below.

(1) Interviewers should record the year of birth **only** when it is reported by the respondent. Interviewer training currently emphasizes the importance of obtaining accurate information on the year of birth. This type of information is, of course, crucial in determining which children should be included in the fertility planning and health and breastfeeding sections of the core questionnaire. The emphasis on the year of birth is also based on the underlying assumption that data obtained in the field by interviewers is more accurate than data imputed by computer in the central office. The latter rationale is less compelling in light of the analysis of data from DHS surveys in Africa. Moreover, a study of age misreporting by the National Academy of Sciences concludes that interviewers should not be encouraged to perform arithmetic in the field because of the possibility of introducing additional errors (Ewbank, 1981). Instead, interviewers should be instructed to record the year of birth **only** when that information is supplied directly by the respondent. The only time that an unusual effort should be made to determine the year of birth in the field (when the respondent cannot initially provide it) is when the age of the child is around 4-6 years, the cut-off point for being included in certain sections of the questionnaire. Alternatively, the questionnaire could be modified so that the fertility planning and health questions are asked either of children born after January 1 of the fifth year before the interview **or** of children age 0-6 years at the time of the survey.

(2) Interviewer training should emphasize how to correctly calculate the year of birth from a child's age and month of birth. Although interviewers in some countries have been carefully trained in methods of calculating year of birth from age and month of birth, it appears that they are not using this procedure correctly when the respondent can give neither the year nor the month of birth. One way to clarify the correct method of age calculation would be to supply interviewers with a conversion chart which can be used to calculate the year of birth directly from a child's reported age. Such a chart would also identify inconsistencies in responses with respect to age and year and month of birth and would allow interviewers to query respondents about such inconsistencies during the interview. Age conversion charts have been used successfully in many countries (including some DHS countries), although they do add some complexity to the interviewer's job. If this procedure is not applied universally, it should at least be considered in African countries where age reporting is likely to be problematic.

(3) The imputation program can be changed to correct the problem of incorrect calculation of the year of birth. The ISSA editing and imputation program has been modified on an ad hoc basis in a few countries in which the standard program would greatly distort the month of birth distribution. Two types of procedures have been used. In Ghana, when age plus the year of birth were equal to the year of the interview and the month of birth needed to be imputed, the recorded year of birth was ignored when imputing the month of birth. The year of birth was then recalculated on the basis of age and the imputed month of birth. In Mali, whenever age and year of birth were reported but the month of birth was missing, recorded age was ignored in imputing approximately half of the cases and recorded year of birth was ignored in the other cases. An imputation procedure that ignores age will not correct the most common type of problem discussed here, whereas a procedure that ignores the recorded year of birth should lead to an accurate result in such cases. Further investigation is needed to determine the types of modifications in the editing and imputation program that are most appropriate. Consideration should also be given to standardizing the changes across countries (particularly in Africa).

(4) A follow-up question after each age question can be used to verify the reported age. In some countries (such as Thailand), many respondents in previous surveys have been found to report age at next birthday rather than age at last birthday, particularly if it is less than six months until the next birthday. Ironically, this error partially offsets the year of birth problem discussed above. Nevertheless, an attempt should be made to determine what age concept respondents are using. The best way to accomplish this is to ask an explicit follow-up question after each age question to determine whether the stated age is the age at the last birthday or the age at the next birthday. These follow-up questions would only have to be used in countries in which substantial numbers of respondents tend to report age at the next birthday for themselves or for their children.

6 Coverage of Live Births

A substantial effort has been made in the DHS surveys to obtain a complete count of all live births of each respondent. Since respondents in past surveys have often forgotten to mention children who have died, children who are not living at home, or children who are very young, special precautions were taken to fully enumerate such children. Respondents were asked separate questions about the number of sons and daughters living with them, the number living elsewhere, and the number who had died. The distinction of children by sex improves the accuracy of reporting. A follow-up question was also asked to ascertain whether the total number of children enumerated was correct. If there are any discrepancies, the interviewer is instructed to probe and to correct the information given previously.

In contrast to many earlier surveys that contained only a question on the date of the last one or two live births (such as the Contraceptive Prevalence Surveys) or a truncated birth history, DHS surveys contain a complete and detailed birth history. The full birth history permits a more thorough assessment of the quality of the data, and the birth data can be compared with those from other surveys for overlapping periods. At the end of the birth history section of the questionnaire, there is a final check to insure that the number of births in the birth history is equal to the sum of sons and daughters ever born (obtained earlier). If there is a discrepancy, the numbers must be reconciled.

In the interviewer's manual (Institute for Resource Development, 1987), interviewers are reminded of the importance of the questions on reproduction and are told to be especially careful to obtain all the required information. Instructions are given to include all live births even if the child survived for only a few minutes. When the interviewer finishes the birth history, she is instructed to ask the respondent explicitly about any periods of more than three years between births or more than three years since the last live birth to see if any births may have been omitted. Interviewer training sessions also emphasize the importance of collecting complete and accurate information about all live births. Although these procedures do not insure that no birth will be missed, they do provide assurance that every possible effort will be made to avoid underreporting of births.

The omission of births is difficult to detect from the data themselves unless there is gross underreporting. The apparent deficit of births in a particular age group may be due either to missing births or to erroneous reporting of birth dates. Moreover, it is hard to distinguish these two types of errors from one another. Typically, when births are omitted, undercoverage is larger for older women, many of whose births occurred further in the past. One way of detecting any large-scale omission of births is to examine the average number of children ever born by the age of respondents. Unless fertility has been rising over time, the average parity of women should be higher at each older age group.

As shown in Table 6.1, which presents the average number of children ever born for each 5-year age group of women, the average parity increases monotonically with age in all 22 DHS countries. Therefore, there is no indication of gross underreporting of births according to this measure. Mali is the only country without an increase of at least 0.4 children in the average parity from age 40-44 to 45-49 years. In contrast, in the World Fertility Survey 6 of the 40 countries for which data were available up to age 45-49 did not exhibit an increase in average parity from age 40-44 to 45-49 and the average parity actually declined in three of those countries (Goldman et al., 1985: Table 15). The average difference in parity at these ages was 0.7 children in the DHS surveys compared with 0.3 children in the WFS surveys. For the 13 countries included in both surveys, the average parity difference between the two oldest age groups was 0.4 children in WFS surveys and twice as large in DHS surveys (0.8 children). This might be taken as an indication of the relatively complete reporting of births among older women in the Demographic and Health Surveys.

When the average number of children ever born from the 13 DHS surveys was reconstructed for the time of the WFS fieldwork (data not shown), the average number of children ever born calculated from the two surveys was generally comparable. This result is encouraging since these births occurred closer to the time of the WFS surveys than to the DHS surveys and since WFS itself had achieved a significant improvement in the coverage of births over censuses and previous surveys (Goldman et al., 1985).

6.1 AGE AT FIRST BIRTH

An examination of the median age at first birth for different age groups of respondents can also help to identify problems in the coverage or reported timing of births (Casterline and Trussell, 1980; Chidambaram et al., 1980; Goldman et al., 1985). If the age at first birth has not changed appreciably over time, then the median age at

Table 6.1 Average number of children ever born by age of woman, Demographic and Health Surveys, 1986-1989

Country	Age of Woman						
	15-19	20-24	25-29	30-34	35-39	40-44	45-49
SUB-SAHARAN AFRICA							
Botswana	0.3	1.2	2.5	3.7	5.1	5.4	5.8
Burundi	a	0.9	2.7	4.2	5.6	6.6	7.3
Ghana	0.2	1.3	2.7	4.2	5.5	6.6	7.3
Liberia	0.5	1.8	3.2	4.2	5.3	5.9	6.8
Mali	0.6	1.9	3.4	5.0	6.0	7.0	7.1
Senegal	0.3	1.6	3.1	4.7	6.2	6.8	7.3
Togo	0.3	1.4	2.9	4.6	5.7	6.9	7.3
Uganda	0.4	1.9	3.6	5.0	6.8	7.2	7.8
Zimbabwe	0.2	1.3	2.9	4.3	5.5	6.4	6.9
NORTH AFRICA							
Morocco	0.1	0.7	2.1	3.8	5.2	6.7	7.3
Tunisia	a	0.5	1.7	3.4	4.7	5.5	6.3
ASIA							
Indonesia	0.1	1.0	2.2	3.4	4.3	5.2	5.6
Sri Lanka	a	0.5	1.4	2.4	3.0	4.0	4.9
Thailand	0.1	0.6	1.4	2.2	3.0	3.9	5.0
LATIN AMERICA/CARIBBEAN							
Brazil	0.1	0.9	1.9	2.9	3.8	4.6	NA
Colombia	0.1	0.9	1.9	2.8	3.8	4.9	6.1
Dominican Republic	0.2	1.1	2.2	3.4	4.6	5.5	7.0
Ecuador	0.2	1.0	2.4	3.5	4.5	5.4	6.0
Guatemala	0.3	1.5	3.0	4.3	5.2	5.9	NA
Mexico	0.2	1.1	2.3	3.5	4.3	5.4	6.4
Peru	0.1	0.9	2.2	3.5	4.9	5.7	6.3
Trinidad and Tobago	0.1	0.9	1.9	2.7	3.2	3.9	4.9

NA = not available
[a] Less than 0.05

first birth should be constant for different cohorts of women. An increase in the age at marriage and the age at first birth over time should be reflected in higher median values for the younger cohorts. For the older cohorts, however, one would expect little variation in the median age at first birth since the first births were taking place about 15 to 30 years before the survey.

In 9 out of 11 DHS surveys in Africa, the median age at first birth is substantially greater for women 45-49 than for women 40-44 (see Table 6.2). This is not the case in Latin America or Asia (with the exception of Ecuador and Indonesia). The results for Africa are probably due either to underreporting of early births among the oldest women (which is especially likely for children who did not live long) or to misreporting of the year of birth or the age of the first child. The latter explanation is consistent with the Potter effect, which is produced when older women report their early children as being born closer to the time of the survey than actually occurred (Potter, 1977).

6.2 SEX RATIOS AT BIRTH

Reported sex ratios at birth that are outside the range usually considered normal could indicate the sex-selective omission of births. Moreover, an examination of the sex ratios at birth for different time periods can reveal changing patterns over time that could also suggest the selective omission of either males or females. The most common pattern would be one of increasing sex ratios for periods further in the past, reflecting an omission of female births (and particularly female deaths) in the earlier periods.

Sex ratios at birth for five-year periods before the survey are shown in Table 6.3. For all births, the sex ratio falls in the range of 103.2-106.9 in 16 of the 22 countries. The overall sex ratio at birth is unusually low (less than 100) in Botswana and Uganda and unusually high (more than 108) in Colombia, Ecuador, and Peru. This would seem to suggest that some male births were omitted in parts of Africa and some female births were omitted in parts of

Table 6.2 Median age at first birth by age of woman at the time of the survey, Demographic and Health Surveys, 1986-1989

Country	Age of Woman					
	20-24	25-29	30-34	35-39	40-44	45-49
SUB-SAHARAN AFRICA						
Botswana	19.7	19.2	19.3	19.6	20.0	20.9
Burundi	a	20.9	21.1	21.1	21.2	21.1
Ghana	19.9	20.0	19.2	19.5	18.8	19.3
Liberia	18.5	19.0	19.4	19.8	18.6	21.0
Mali	18.4	19.0	18.6	19.1	18.7	20.3
Senegal	19.0	19.0	19.0	18.7	18.7	19.2
Togo	19.5	19.2	18.8	19.5	19.2	20.0
Uganda	18.6	18.3	18.0	18.0	18.3	18.6
Zimbabwe	a	19.5	19.4	19.8	19.2	19.7
NORTH AFRICA						
Morocco	a	22.4	21.3	21.0	20.1	20.0
Tunisia	a	24.5	23.2	22.5	21.5	22.4
ASIA						
Indonesia	a	20.2	19.9	19.6	19.3	19.8
Sri Lanka	a	24.7	24.1	24.9	23.1	21.9
Thailand	a	23.0	22.7	22.3	22.2	21.6
LATIN AMERICA/CARIBBEAN						
Brazil	a	22.4	22.8	22.4	22.2	NA
Colombia	a	21.6	21.9	21.6	20.8	21.0
Dominican Republic	a	21.2	20.5	19.9	19.9	19.7
Ecuador	a	20.7	21.0	21.2	21.1	21.6
Guatemala	20.0	19.7	19.7	20.1	20.2	NA
Mexico	a	21.1	20.6	21.5	21.0	20.7
Peru	a	21.4	21.4	21.2	20.8	20.5
Trinidad and Tobago	a	22.2	21.9	21.6	21.1	20.5

NA = Not available

[a] Fewer than half of the women in the age group have given birth by age 20.

Latin America. For births that occurred at least 25 years before the survey, the sex ratio is high in Burundi (126.7), Mali (121.7), and the Dominican Republic (119.4).

Although such high sex ratios suggest a selective omission of female births during that period, it should be noted that they are based on relatively small numbers of births and the sampling errors of the estimates are large. Moreover, in the past, lower parity has been found to be associated with higher sex ratios at birth, although the effects tend to be small (Markle, 1974; Teitelbaum, 1972). Therefore, it is not possible to conclude, on the basis of the sex ratio estimates alone, that there was gross omission of female births in the past in any of these countries.

Table 6.3 Sex ratios at birth for five-year periods prior to the survey, Demographic and Health Surveys, 1986-1989

	Years Prior to the Survey						
	0-4	5-9	10-14	15-19	20-24	25+	Total
SUB-SAHARAN AFRICA							
Botswana	97.6	97.1	96.5	100.5	112.0	113.9	99.0
Burundi	104.9	99.5	100.1	123.6	106.8	126.7	105.1
Ghana	103.4	108.0	106.6	103.2	102.4	101.6	105.0
Liberia	106.1	101.9	106.8	109.7	96.3	101.3	104.6
Mali	105.4	103.8	107.8	111.2	106.4	121.7	106.9
Senegal	103.6	106.8	101.0	107.0	105.6	103.3	104.5
Togo	102.7	107.9	101.3	98.3	99.3	88.3	102.3
Uganda	99.5	102.6	94.0	100.1	101.7	106.9	99.8
Zimbabwe	101.1	101.2	110.0	105.2	101.6	103.3	103.5
NORTH AFRICA							
Morocco	100.3	107.7	104.4	103.4	108.2	99.9	104.1
Tunisia	101.2	104.9	102.8	99.9	108.8	115.5	103.4
ASIA							
Indonesia	109.9	104.9	102.8	102.2	106.6	112.1	105.7
Sri Lanka	111.9	105.5	105.1	102.6	100.8	107.1	105.9
Thailand	106.7	105.3	107.5	110.4	101.2	101.9	106.3
LATIN AMERICA/CARIBBEAN							
Brazil	107.0	100.8	104.9	108.7	109.1	108.5	105.2
Colombia	105.6	109.4	110.9	107.0	112.3	107.0	108.6
Dominican Republic	98.9	107.7	108.1	100.9	95.1	119.4	103.7
Ecuador	109.3	102.2	107.2	108.9	131.2	112.2	108.7
Guatemala	99.3	108.1	105.4	119.0	111.8	(150.0)[a]	106.4
Mexico	103.2	101.0	107.3	101.8	107.9	92.8	103.2
Peru	105.9	104.6	106.3	114.7	115.2	110.5	108.1
Trinidad and Tobago	99.0	105.9	111.6	104.4	114.6	112.1	106.3

[a]Based on fewer than 100 female births.

7 Comparison of DHS and WFS Fertility Rates

A standard method for evaluating the quality of fertility data is to compare fertility rates from different sources for overlapping periods of time. In most DHS countries, fertility rates are available from censuses or other national-level surveys for earlier periods of time. The retrospective birth histories in DHS surveys also allow fertility rates to be calculated for past periods, although rates for earlier periods can be calculated only for younger women. The final reports for many DHS surveys contain a detailed comparison of fertility estimates from a number of sources. It is not possible in this report to make these comparisons for each of the 22 countries. However, a comparison will be made of DHS and WFS fertility estimates for the 13 countries included in both survey projects.

The comparison is facilitated by the fact that the methodology was similar in both surveys and the samples were usually national in scope. The main difference is that WFS included a complete pregnancy history whereas DHS restricted its fertility questions to live births. Figure 7.1 compares trends in the total fertility rate (up to age 35) in 13 countries, based on data from DHS and WFS surveys. In all cases, the DHS estimates are for the three most recent five-year periods before the survey. Complete information on fertility during that fifteen-year period is available only through age 30-34, so the total fertility estimates are limited to women age 15-34. Each five-year period is centered on a point shown on the graph.

The WFS estimates are taken from a variety of published sources. Since the WFS First Country Reports showed age-specific fertility rates for time periods of different lengths and since some of those reports did not include age-specific fertility rates at all, it is not possible to use a single series of standard time periods for the WFS estimates. Some of the fertility estimates in Figure 7.1 are annual estimates, others are multiple-year estimates, and some are moving averages of annual fertility estimates.

Figure 7.1 shows that agreement between the DHS and WFS fertility estimates for overlapping periods is excellent in most cases. If anything, DHS estimates would be expected to be somewhat lower, since the births under consideration occurred further in the past and women may have forgotten to report them, particularly if the children died at a young age. The DHS and WFS estimates are quite consistent in most cases, however, and when there is a discrepancy, the DHS estimates are always higher. This would suggest that: (1) the reporting of births for the overlapping period was more complete in DHS, (2) WFS births that occurred close to the time of the survey were reported as occurring earlier than they actually did, or (3) older women in the DHS surveys underestimated the age of their older children. In any case, the comparison of DHS and WFS fertility data provides no evidence of a substantial underreporting of births in DHS surveys.

The only country in which the two estimates are not in general agreement is Indonesia (see Figure 7.1). This is explained by the fact that the WFS survey in Indonesia covered only Java and Bali—which are the regions with the lowest fertility levels—while the DHS survey was carried out in 20 of Indonesia's 27 provinces. Most of the discrepancy between the fertility estimates can be attributed to the different geographical areas included in the surveys. The DHS fertility estimate is more representative of Indonesia as a whole since the DHS sample represents 93 percent of the total population.

Figure 7.1 Trends in the total fertility rate (TFR) for women 15-34, selected DHS and WFS surveys

SUB-SAHARAN AFRICA

NORTH AFRICA

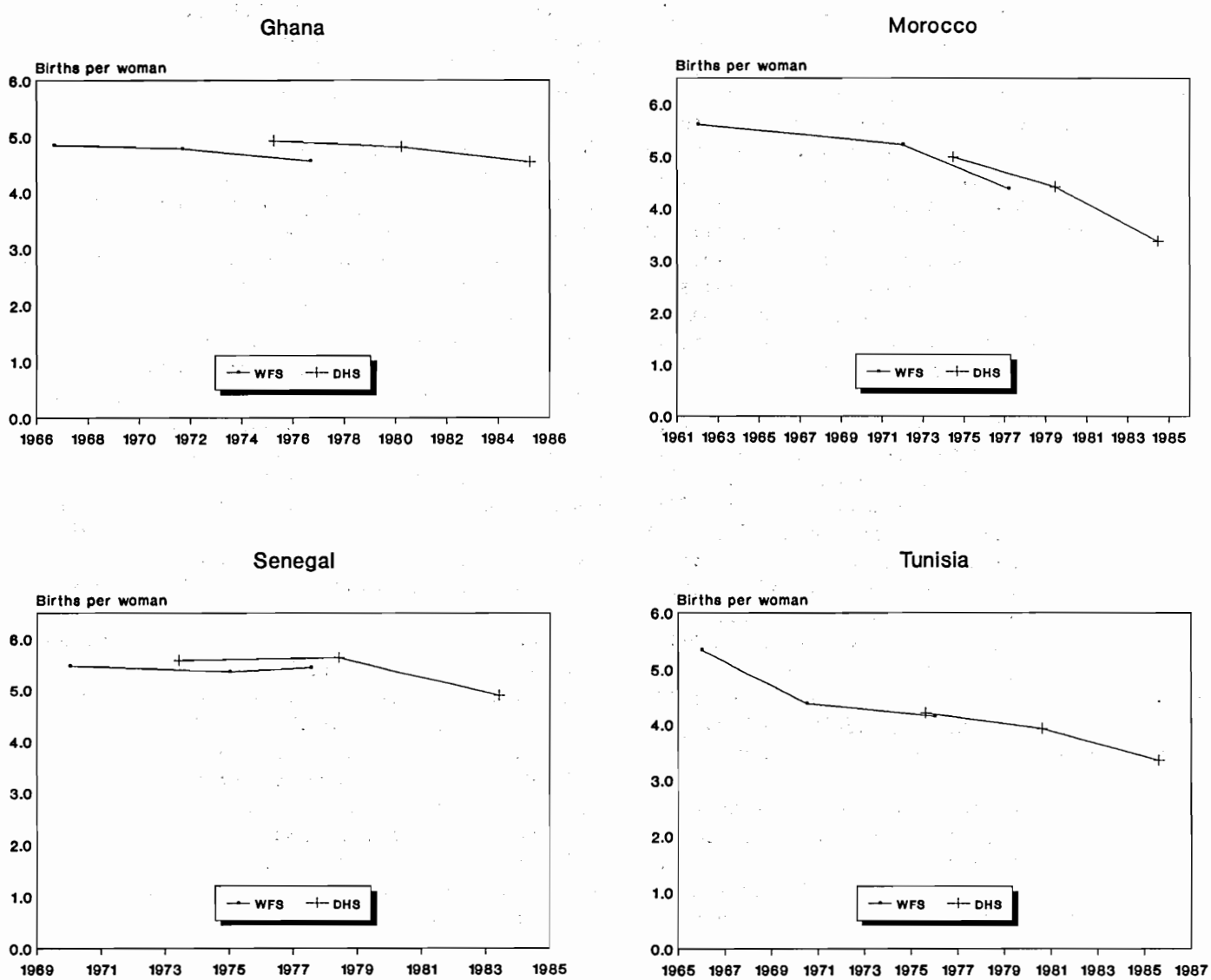
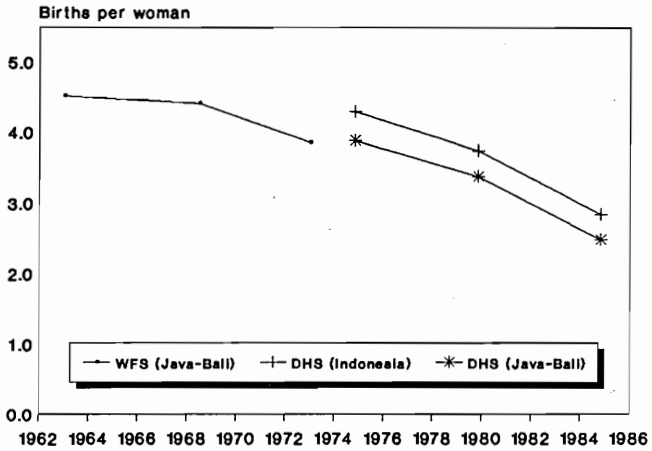


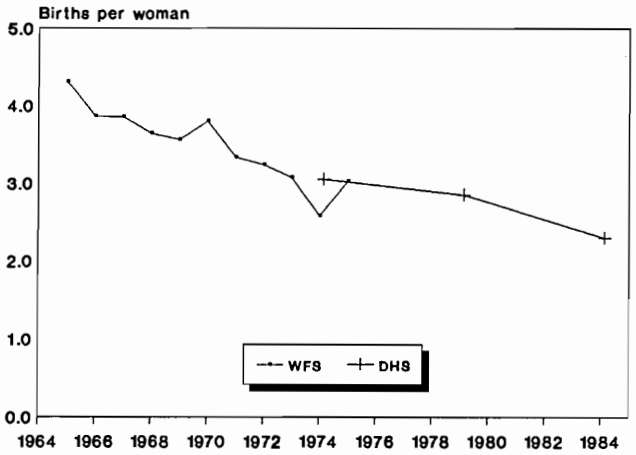
Figure 7.1—*Continued*

ASIA

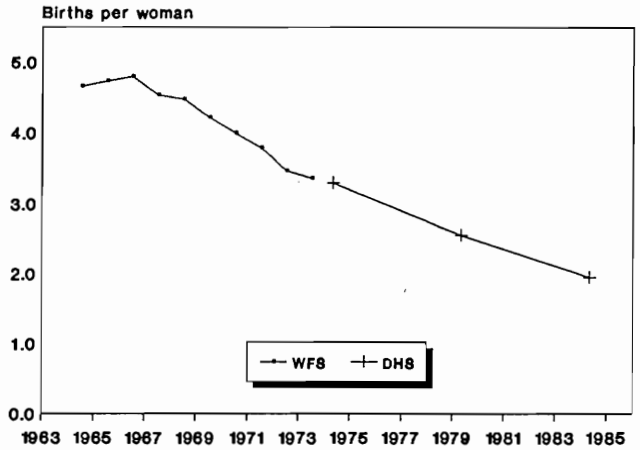
Indonesia



Sri Lanka

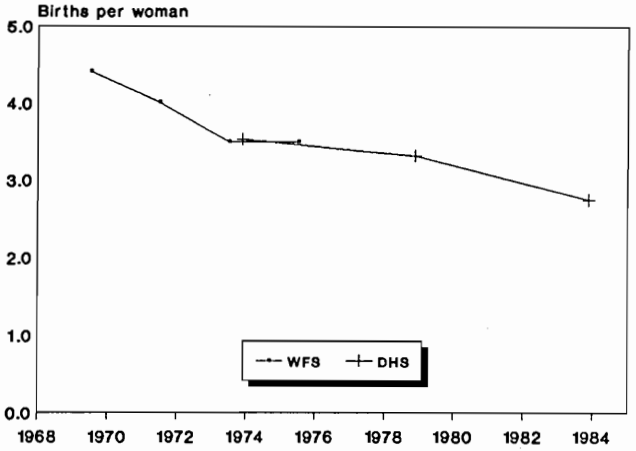


Thailand

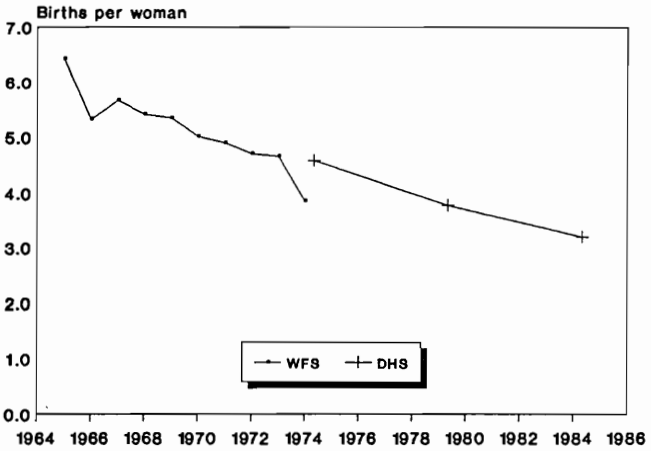


LATIN AMERICA/CARIBBEAN

Colombia



Dominican Republic



Ecuador

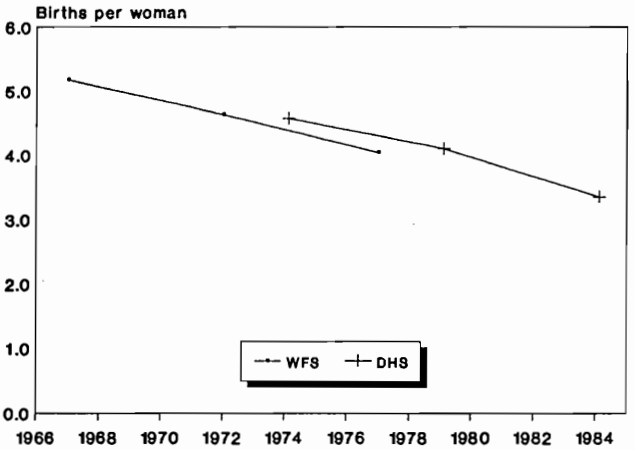
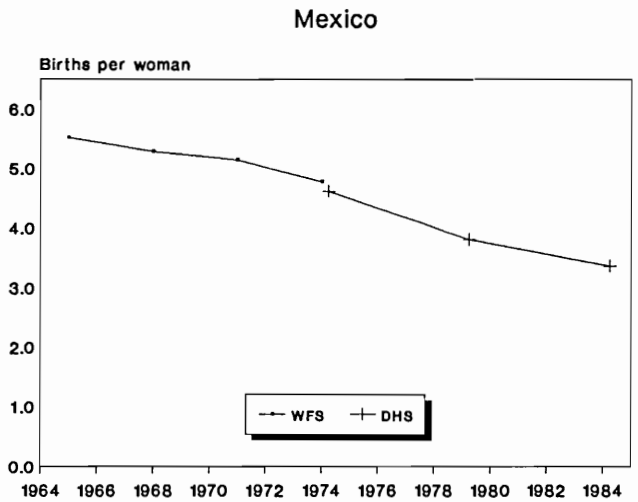
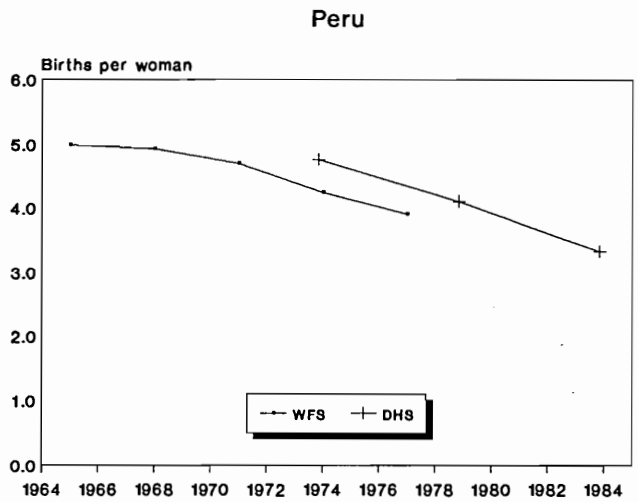
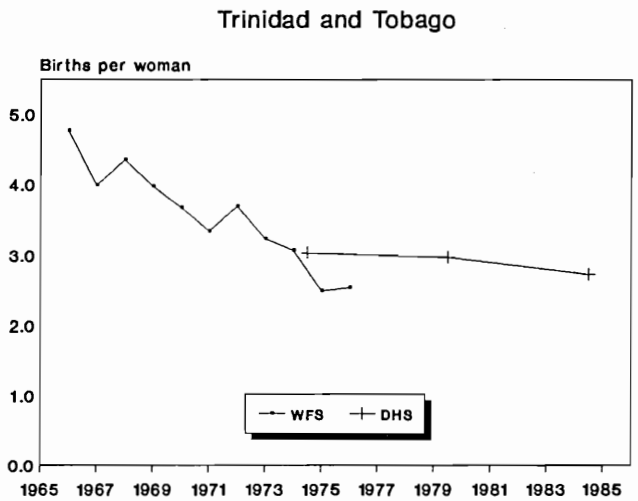


Figure 7.1—*Continued*







Sources for Figure 7.1:

Colombia: Departmento Administrativo Nacional de Estadística (1977)
Dominican Republic: Consejo Nacional de Poblacion y Familia (1976)
Ecuador: Instituto Nacional de Estadística y Censos and World Fertility Survey (1984)
Ghana: Central Bureau of Statistics (1978); United Nations (1987)
Mexico: Dirección General de Estadística (1979)
Morocco: Farid (1984); United Nations (1987)
Peru: Oficina Nacional de Estadística (1979)
Senegal: Direction de la Statistique and World Fertiltiy Survey (1981); Gueye (1984)
Sri Lanka: Retherford and Alam (1985)
Thailand: Piampiti and Knodel (1978); two-year moving average.
Trinidad and Tobago: Central Statistical Office (1981)
Tunisia: United Nations (1987)

8 Summary and Conclusions

This report assesses the quality of DHS birth history data and fertility rates using the results from DHS surveys in 22 countries. Because of the high standards observed in DHS surveys, reasonably complete and accurate information on births has been obtained in all DHS countries. The existence of multiple questions on fertility and internal checks on the data, as well as the careful training and supervision of interviewers, have contributed to the accuracy of the results. Nevertheless, large-scale demographic surveys cannot avoid errors in coverage and timing of births, and methods of improving fertility estimates need to be actively pursued.

In this report, several methods were used to evaluate the completeness of birth reporting in DHS surveys. These include an examination of:

- the average number of children ever born by age,
- fertility data from other sources,
- the sex ratio of births at different time periods, and
- the median age at first birth for different age groups.

The parity analysis did not reveal gross omission of births in any countries and DHS estimates of fertility were found to compare favorably with World Fertility Survey estimates for overlapping time periods in the same countries. Two measures, the sex ratio of births and the median age at first birth, however, exhibited moderate distortions in the pattern of results for some countries, primarily in Africa.

In addition, problems were detected in the reporting of children's dates of birth and ages, and these errors may affect estimates of current fertility and trends in fertility over time. Four specific problems were identified: (1) the displacement of children's birth dates to avoid asking a series of questions on fertility planning, health, and breastfeeding, (2) miscalculation of the year of birth, (3) age heaping or digit preference, and (4) missing or incomplete information on some birth dates. Despite the existence of several reporting problems, their impact is attenuated by the fact that fertility rates are generally reported for multi-year periods and some births are displaced within a single reference period. The analyses conducted in this study indicate that these problems are not universal and that, in most of the countries where they do occur, their impact on fertility rates is small. Nevertheless, steps should be taken to reduce the incidence of reporting errors and their impact on demographic rates, particularly in Africa, where they are most prevalent.

On the basis of the findings on birth reporting, several recommendations were made for changes in survey design and survey procedures:

(1) Careful training and supervision of interviewers is necessary to reduce the problem of displacement of birth dates. In particular, supervisors should personally recheck (whenever possible) the birth dates of children born in the year or two before the eligibility cut-off date for the health and breastfeeding section of the questionnaire. Interviewers should be made aware of the scrutiny such births will encounter, in order to deter them from recording incorrect information initially.

(2) Consideration should be given to modifying the reference periods used for the calculation of demographic rates. Since the displacement problem particularly affects rates calculated for five-year periods, it might be preferable to utilize three-year rates for estimates of current fertility and mortality in countries where displacement is likely to be a problem. Four-year rates are most appropriate for analyzing demographic trends over time. These procedures would insure that the reference periods for demographic rates are not coincident with the cut-off date for eligibility for the health and breastfeeding questions.

(3) Since displacement occurs most frequently in the case of dead children, it is desirable to reduce the number of questions asked about children who have died.

(4) Interviewers in countries in which there are problems with the reporting of birth date and age should be supplied with a chart for converting age to year of birth. Since interviewers sometimes make errors performing these calculations themselves, the use of age conversion charts will help reduce the magnitude of this type of error and facilitate the recording of accurate and consistent birth date information. Some countries have already used this procedure, with considerable success. Another possibility is to give the interviewers more leeway in recording the year of birth as "unknown" when the respondent appears incapable of supplying the information.

(5) Where the reporting of age, instead of birth date, is common, explicit follow-up questions could be asked to determine whether the respondent is reporting age at the last birthday or age at the next birthday.

(6) Consideration should be given to modifying the ISSA editing and imputation program to avoid bunching imputed birth dates in the early months of the year.

These suggestions should be evaluated not only with respect to their consequences for the estimation of fertility rates, but also with respect to their impact on such things as survey costs and ease of implementation. The possibility also exists that additional problems may be generated in the process of carrying out these recommendations. Moreover, it may not be necessary to implement these changes across the board, since many countries do not exhibit the problems described above. The changes suggested can easily be incorporated in future DHS surveys and may lead to significant improvements in the fertility estimates obtained from national sample surveys.

References

Arnold, Fred and Ann K. Blanc. 1990. *Fertility Levels and Trends*. DHS Comparative Studies, No. 2. Columbia, Maryland: Institute for Resource Development.

Blanc, Ann, Trevor Croft, Kia Reinis, and Naomi Rutenberg. 1989. *DHS-II A-Core Questionnaire Field Trial Report*. Columbia, Maryland: Institute for Resource Development.

Casterline, John R. and James Trussell. 1980. *Age at First Birth*. WFS Comparative Studies, No. 15. Voorburg, Netherlands: International Statistical Institute.

Central Bureau of Statistics [Ghana]. 1983. *Ghana Fertility Survey, 1979-1980: First Report*. Accra: Central Bureau of Statistics.

Central Bureau of Statistics [Indonesia]. 1978. *Indonesia Fertility Survey 1976: Principal Report*. Jakarta: Central Bureau of Statistics.

Central Statistical Office [Trinidad and Tobago]. 1981. *Trinidad and Tobago Fertility Survey: 1977*. Port of Spain: Central Statistical Office.

Chidambaram, V.C. and Zeba A. Sathar. 1984. *Age and Date Reporting*. WFS Comparative Studies, No. 5. Voorburg, Netherlands: International Statistical Institute.

Chidambaram, V.C., John G. Cleland, and Vijay Verma. 1980. *Some Aspects of WFS Data Quality: A Preliminary Assessment*. WFS Comparative Studies, No. 16. Voorburg, Netherlands: International Statistical Institute.

Consejo Nacional de Población y Familia [Dominican Republic]. 1976. *Encuesta Nacional de Fecundidad: Informe General*. Santo Domingo: Consejo Nacional de Población y Familia, Secretaría de Estado de Salud Pública y Asistencia Social, República Dominicana.

Departmento Administrativo Nacional de Estadística [Colombia]. 1977. *Encuesta Nacional de Fecundidad: Colombia 1976*. Bogota: Instituto Internacional de Estadística, Corporación Centro Regional de Población.

Dirección General de Estadística [Mexico]. 1979. *Encuesta Mexicana de Fecundidad: Primer Informe Nacional*. Mexico City: Secretaría de Programación y Presupuesto, Coordinación General del Sistema Nacional de Información, Dirección General de Estadística.

Direction de la Statistique [Senegal] and World Fertility Study. 1981. *Enquête Sénégalaise sur la Fécondité, 1978: Rapport National d'Analyse*. Dakar: Division des Enquêtes et de la Démographie, Direction de la Statistique.

Ewbank, Douglas C. 1981. *Age Misreporting and Age-Selective Underenumeration: Sources, Patterns, and Consequences for Demographic Analysis*. Committee on Population and Demography, Report No. 4. Washington, D.C.: National Academy Press.

Farid, Samir. 1984. Fertility Patterns in the Arab Region. *International Family Planning Perspectives* 10(4):119-25.

Goldman, Noreen, Shea O. Rutstein, and Susheela Singh. 1985. *Assessment of the Quality of Data in 41 WFS Surveys: A Comparative Approach*. WFS Comparative Studies, No. 44. Voorburg, Netherlands: International Statistical Institute.

Gueye, Lamine. 1984. *Enquête Sénégalaise sur la Fécondité: Rapport d'Evaluation*. WFS Scientific Reports, No. 49. Voorburg, Netherlands: International Statistical Institute.

Institute for Resource Development (IRD). 1987. *Interviewer's Manual for Use with Model "A" Questionnaire for High Contraceptive Prevalence Countries*. Basic Documentation, No. 5. Columbia, Maryland: IRD.

Institute for Resource Development (IRD). 1989. *Demographic and Health Surveys Data Processing Manual*. Columbia, Maryland: IRD.

Instituto Nacional de Estadística y Censos [Ecuador] and World Fertility Survey. 1984. *Encuesta Nacional de Fecundidad: Ecuador-1979*. Quito: Instituto Nacional de Estadística y Censos.

Markle, G.E. 1974. Sex Ratio at Birth: Values, Variance, and Some Determinants. *Demography* 11(1):131-142.

Nagi, Moni H., E.G. Stockwell, and L.M. Snavley. 1973. Digit Preference and Avoidance in the Age Statistics of Some Recent African Censuses: Some Patterns and Correlates. *International Statistical Review* 41(2):165-174.

Oficina Nacional de Estadística [Peru]. 1979. *Encuesta Nacional de Fecundidad del Perú: 1977-1978*. Lima: Dirección de Demografía, Dirección General de Censos, Encuestas y Demografía.

Piampiti, Sauvaluck and John Knodel. 1978. Revised Estimates of Age-Specific Fertility Rates from the Survey of Fertility in Thailand. *World Fertility Survey Supplement,* Vol. 1, No. 1. Bangkok, Thailand: Institute of Population Studies, Chulalongkorn University and National Statistical Office.

Potter, Joseph E. 1977. Problems in Using Birth History Analysis to Estimate Trends in Fertility. *Population Studies* 31(2):335-364.

Retherford, Robert D. and Iqbal Alam. 1985. Comparison of Fertility Trends Estimated Alternatively for Birth Histories and Own Children. *Papers of the East-West Population Institute*, No. 94. Honolulu, Hawaii: East-West Center.

Scott, Christopher and G. Sabagh. 1970. The Historical Calendar as a Method of Estimating Age: The Experience of the Moroccan Multi-Purpose Survey of 1961-63. *Population Studies* 24(1):93-109.

Sepulveda, Bienvenida R. 1984. *Evaluación de la Encuesta Nacional de Fecundidad de la República Dominicana de 1980.* WFS Scientific Reports, No. 63. Voorburg, Netherlands: International Statistical Institute.

Stockwell, E.G. 1966. Patterns of Digit Preference and Avoidance in the Age Statistics of Some Recent Nation l Censuses: A Test of the Turner Hypothesis. *Eugeni Quarterly* 13(3):205-208.

Teitelbaum, Michael S. 1972. Factors Associated with the Sex Ratio in Human Populations. In *The Structure of Human Populations*, ed. G.A. Harrison and A.J. Boyce, 90-109. London: Oxford University Press.

Turner, Stanley H. 1958. Patterns of Heaping in the Reporting of Numerical Data. Proceedings of the Social Statistics Section. Washington, D.C.: American Statistical Association.

United Nations. 1987. *A Comparative Evaluation of Data Quality in Thirty-Eight World Fertility Surveys*. New York: United Nations, Department of International Economic and Social Affairs.

United Nations. 1983. *Fertility Levels and Trends as Assessed from Twenty World Fertility Surveys*. New York: United Nations, Department of International Economic and Social Affairs.

Weller, Robert H., William J. Serow, and Mohamed Bailey. 1987. Palestinian Refugee Statistics, Digit Preference and the Turner Hypothesis. *Population Bulletin of the ESCWA* 31:5-23.

Assessment of the Quality of Data Used for the Direct Estimation of Infant and Child Mortality in the Demographic and Health Surveys

JEREMIAH M. SULLIVAN
GEORGE T. BICEGO
SHEA OSCAR RUTSTEIN

Acknowledgments

The authors wish to acknowledge Stanley Becker and Anne Cross for their insights and suggestions on earlier drafts of this report, and Guillermo Rojas for his help in programming.

1 Introduction

This report assesses the quality of the Demographic and Health Surveys (DHS) data used for the direct estimation of infant and child mortality rates. The assessment covers data from 22 surveys in developing countries (11 in Africa, 3 in Asia, and 8 in Latin America and the Caribbean). Mortality levels vary greatly, with estimates of infant mortality ranging from 25 per 1000 live births (Sri Lanka and Trinidad and Tobago) to 130 per 1000 (Liberia and Mali).

The primary objective of this analysis is to identify errors in data collection which have occurred in a number of surveys and which may signal a need to modify the DHS questionnaires or field procedures. A second objective is to provide users of DHS data with comprehensive, as well as survey-specific, information about the quality of the data used to calculate childhood mortality rates. The report is designed as a comparative analysis, with the same tests of data quality applied to each survey. This approach was used previously to evaluate the mortality data from the World Fertility Survey (WFS) (Rutstein, 1985).

Direct calculation of infant and child mortality rates requires data from survey respondents regarding children's date of birth, survival status, and age at death for those who died. While mortality rates have associated sampling errors from these data, only nonsampling error is considered in this analysis. Nonsampling error occurs most frequently as a result of underreporting of live births (especially children who died), or due to inaccurate information on date of birth or age at death. These sources of error are considered below. Section 2 describes the DHS data collection procedures. Section 3 considers the date of birth data and Section 4 the age at death data. Event underreporting is discussed in Section 5, including the results of tests of internal consistency and comparison of DHS mortality estimates with estimates from other sources. Section 6 summarizes results.

2 Collection of DHS Mortality Data

2.1 QUESTIONNAIRES

The data for calculating mortality rates were collected in the section on reproduction in the individual questionnaire (see Appendix page 139). The respondent is asked to report all live births, including births of children who have died, first in terms of the aggregate number of children ever born (CEB), then in terms of specific questions about each live birth (birth history).

The CEB data are collected in a series of six questions: the number of sons and daughters who are living at home, the number of sons and daughters living elsewhere, and the number of each who have died. If a respondent reports that she has had no children who have died, she is asked if she ever gave birth to a baby that cried or showed movement but is no longer alive.

The birth history data are collected in chronological order, starting with the first birth and ending with the most recent. Data are collected on the date of birth, sex, survivorship status, current age, and whether living with the mother (for living children) or age at death (for dead children). Following the birth history, the interviewer checks the consistency of the CEB and the birth history data and reconciles differences.

Specific rules are applied to the collection of data on date of birth (year and month), current age, and age at death. In the case of **year** of birth, current age, and age at death, the rules state that a date or age must be recorded (i.e., the code for "don't know" is not acceptable). If a respondent could not readily provide this information, probing techniques were used. As a last resort, interviewers were instructed to use whatever information was available (e.g., birth dates reported for other children, number of years ago that the birth occurred, etc.) to estimate **year** of birth, current age, and age at death. In the case of **month** of birth the rules for data collection are less stringent. Interviewers are instructed to utilize whatever information is available to determine month of birth (e.g., season), but, when it is not possible to make a reliable estimate, it is acceptable to record the code for "don't know".

Interviewers are trained to probe birth intervals longer than three years in length in order to detect unreported births. If the respondent could not provide a plausible explanation for a long interval between births (e.g. husband away, contraception, etc.), the interviewer was instructed to ask about births which occurred during the interval and may have survived only a short time.

2.2 FIELD PROCEDURES

In the DHS surveys, interviewer training generally lasts three to four weeks. Particular emphasis is placed on the section of the individual questionnaire on reproduction because it is the source of data for the direct estimation of fertility and mortality and also because it identifies the children about whom the health questions are asked later in the questionnaire.

For the same reasons, field editing focuses particularly on the data collected in the birth history. Questionnaires are edited while the interview team is working in a sampling point, so that households can be revisited, if necessary. Field editors are responsible for checking the completeness and consistency of the birth history data. While standard procedures for conducting these checks are incorporated in the instruction manuals and training for each survey, the degree to which these procedures were followed, no doubt, varied between surveys.

2.3 DATA PROCESSING

Data from the survey questionnaires are transformed into standard recode data files following prescribed rules for data entry and editing, date imputation, and recode file creation (see Institute for Resource Development, 1989). Period mortality rates presented here and in the final survey reports are calculated from the recode data files using a method originally developed for the World Fertility Survey (see Rutstein, 1985). The rates shown are the probability of dying between two exact ages: birth and age one for infant mortality, age one and age five for child mortality, and birth and age five for under-five mortality.

3 Date of Birth Data

Date of birth data are essential for any analysis of mortality by time period. When the year or month of birth was not reported, the missing information was imputed. Using other information reported by the respondent, the imputation procedure established a logical time period within which the birth probably occurred and then randomly assigned a date within that time period. Thus, in the standard recode data files, all births have a year and month of birth and each birth is flagged by a variable which indicates the extent of imputation and the type of information originally recorded on the questionnaire.

3.1 COMPLETENESS OF THE DATA

For each DHS survey, the percentage of births with incomplete information on date of birth is shown in Table 3.1. The percentages are for births with a date of birth (reported or imputed) during the fifteen-year period preceding the survey. Statistics are shown separately for living and dead children.

For living children, the percentage of births with missing data ranges from a high of 58 and 39 percent in Mali and

Table 3.1 Percentage of births during the fifteen-year calendar period preceding the survey with incomplete information on date of birth by survival status, Demographic and Health Surveys, 1986-1989

	Living Children			Dead Children			
Country	Any Information Missing	Month Only Missing	Yr., Mo. and Age Missing	Any Information Missing	Month Only Missing	Year and Month Missing	All Births
SUB-SAHARAN AFRICA							
Botswana	2.3	2.0	0.3	7.5	4.0	3.5	2.6
Burundi	14.4	14.4	0.0	35.4	34.8	0.6	18.1
Ghana	17.5	17.5	0.0	36.7	35.5	1.2	24.7
Liberia	11.1	11.1	0.0	25.4	20.9	4.5	14.3
Mali	58.2	58.1	0.1	71.7	45.3	26.4	61.8
Senegal	15.3	15.2	0.1	33.3	31.5	1.8	19.0
Togo	39.0	39.0	0.0	68.3	63.3	5.0	43.7
Uganda	0.0	0.0	0.0	0.2	0.2	0.0	0.1
Zimbabwe	0.2	0.2	0.0	2.7	1.8	0.9	0.4
NORTH AFRICA							
Morocco[a]	32.1	32.1	0.0	65.9	63.3	2.6	36.3
Tunisia	1.9	1.9	0.0	27.7	16.3	11.4	3.9
ASIA							
Indonesia[a]	14.4	14.4	0.0	38.9	38.7	0.2	17.2
Sri Lanka[a]	4.4	4.4	0.0	21.7	14.2	7.5	5.3
Thailand	3.4	3.0	0.4	53.8	23.1	30.7	6.2
LATIN AMERICA/CARIBBEAN							
Brazil	0.7	0.6	0.1	20.5	13.1	7.4	2.7
Colombia	0.9	0.9	0.0	10.7	10.7	0.0	1.5
Dominican Republic	0.9	0.9	0.0	7.5	7.5	0.0	1.5
Ecuador	2.7	2.7	0.0	19.4	14.2	5.2	4.3
Guatemala	1.3	1.3	0.0	16.1	15.1	1.0	3.1
Mexico	0.5	0.5	0.0	7.2	5.7	1.5	1.1
Peru	0.7	0.7	0.0	8.9	8.9	0.0	1.7
Trinidad & Tobago	0.1	0.1	0.0	8.3	6.3	2.0	0.4

Note: Based on births with a reported or imputed date of birth in the fifteen-year calendar period.

[a]An acceptable response for the DHS survey in Morocco included "season" of birth. The month-missing category includes those births for which a season was reported. A season, however, requires imputation within only a 3-month window as opposed to a 12-month window.

Togo, respectively, to a low of 1 percent in Trinidad and Tobago, Uganda, and Zimbabwe. Overall, the majority of these cases are missing only the month of birth, so that imputation is over a range of less than twelve months.[1] The percentage of cases for which the range of imputation was greater than one year (i.e., missing both year of birth and age) was less than one percent in all surveys.

For dead children, the percentage missing data on date of birth is greater than for living children, ranging from a high of 72 and 68 percent in Mali and Togo, to a low of 3 and 1 percent in Zimbabwe and Uganda, respectively. For most surveys, the majority of cases are missing only the month of birth; for example, in Burundi, Ghana, and Senegal information is missing for about 30 percent of dead children, but less than 2 percent are missing year of birth.

However, there are a number of surveys in which a significant percentage of dead children are missing information on year of birth. In seven surveys, 5 percent or more of dead children were missing the year of birth—Brazil, Ecuador, Mali, Sri Lanka, Thailand, Togo, and Tunisia. Mali and Thailand stand out particularly, with 26 and 31 percent of cases (respectively) missing data on year of birth.

3.2 ACCURACY OF THE DATA

The accuracy of DHS data on date of birth is examined in a report by Arnold in this volume. That analysis focuses on a problem found in many DHS surveys: a deficit of births in the fifth calendar year before the survey and an excess in the sixth calendar year. The problem appears to be related to the significant number of health questions which are asked for each birth occurring after the cutoff date—usually January 1 of the fifth calendar year before the survey. Interviewers interested in reducing their workload may have displaced births from the fifth to the sixth calendar year before the survey in order to avoid having to ask health questions about the child.

Birth displacement can affect mortality estimates when the displacement occurs between the reference periods for which the rates are calculated.[2] In DHS survey reports and in this report, mortality rates are calculated for five-year calendar periods. Displacement of births from the fifth to the sixth calendar year moves births and some deaths (depending on the age at death) out of the most recent five-year reference period and into the earlier period. The potential for bias depends on the level of displacement and whether or not it is related to survivorship.

In the report by Arnold (in this volume) an index of birth displacement has been computed as the ratio of births in the fifth calendar year before a survey to half the sum of births in the fourth and sixth calendar years multiplied by 100. Assuming that the actual number of births changes linearly over these three years, the expected value of this ratio is 100. A value less than 100 implies fewer births than expected in the fifth calendar year. Table 3.2 shows the values of birth ratios for the 13 surveys in which the age ratio differs from 100 by 5 points or more. Ten surveys have ratios below 95, suggesting displacement of births from the fifth to the sixth calendar year preceding the survey.

3.3 IMPACT ON ESTIMATES OF INFANT MORTALITY

Birth ratios for living and dead children are shown separately in Table 3.2. In twelve surveys, birth ratios are higher for living than for dead children. This type of birth displacement structure tends to cause negative bias in estimates of infant mortality for the recent period and positive bias in estimates for the preceding period. In ten of these surveys the difference between birth ratios is in excess of 10 points; in six surveys the difference clusters around 18 points (Ecuador, Guatemala, Mali, Peru, Sri Lanka, and Uganda), in three surveys around 30 points (Botswana, Brazil, and the Dominican Republic), and in one survey is 60 points (Trinidad and Tobago).

A simulation model was developed to show the impact of differential levels of displacement on estimates of infant mortality for the two five-year time periods preceding a survey.[3] The model indicated that a birth ratio which is lower for dead children than for living children by 18 points tends to underestimate the more recent infant mortality rate (IMR) by about 2.5 percent and to overestimate

[1] For these cases, the imputed date of birth is accurate to within one year if the recorded year of birth (or age) is correct. The fact that a year of birth (or age) was recorded does not necessarily mean that the information was accurate, only that, if a respondent could not supply the information, the interviewer was able to estimate the year of birth (or age).

[2] It is also possible that displacement of a birth within a reference period could change the period in which the death occurred and thus affect mortality estimates. However, the probability of this happening is small.

[3] The principal assumptions of the model are that the true distribution of births and deaths is uniform within five-year periods, that departures of the birth ratio from 100 for living and dead children is entirely the result of displacement from the fifth to the sixth calendar year preceding a survey, and that displaced dead children died in infancy. The last assumption is conservative in the sense that it maximizes the impact that differential birth displacement has on period mortality estimates. The model showed that the impact of birth displacement on estimates of infant mortality is primarily determined by differential displacement associated with survivorship status and that factors such as the level of mortality and time trends in mortality are of secondary importance.

the IMR for the preceding period to the same degree. In the case where the birth ratio for dead children is lower by 30 points, IMR estimates tend to be underestimated and overestimated, respectively, for the two time periods by about 4 percent. The model was also used to estimate adjusted IMRs for Trinidad and Tobago, with the result that the adjusted estimates were 28 and 35 per 1000 births, compared with the observed rates of 26 (1982-87) and 37 (1977-81), respectively.

Table 3.2 Birth ratios for the fifth calendar year preceding the survey by survival status, Demographic and Health Surveys, 1986-1989

Country	Living Children	Dead Children	Difference in Birth Ratios[a]	All Births
SUB-SAHARAN AFRICA				
Botswana	95.1	66.7	+28.4	93.4
Burundi	82.5	85.7	-3.2	82.9
Ghana	99.7	107.3	-7.6	100.8
Liberia	70.6	71.2	-0.6	70.7
Mali[b]	91.4	74.8	+16.6	87.3
Senegal[b]	87.7	78.5	+9.2	83.6
Togo[b]	86.8	95.3	-8.5	88.7
Uganda	98.9	84.9	+14.0	96.2
Zimbabwe	95.1	126.0	-30.9	97.2
NORTH AFRICA				
Morocco	91.0	95.7	-4.7	91.5
Tunisia[b]	98.9	105.2	-6.3	99.4
ASIA				
Indonesia	110.0	113.8	-3.8	110.4
Sri Lanka	94.4	75.7	+18.7	93.6
Thailand	95.6	114.3	-18.7	96.5
LATIN AMERICA/CARIBBEAN				
Brazil	99.9	73.2	+26.7	97.3
Colombia	95.7	106.1	-10.4	96.4
Dominican Rep.	103.2	72.0	+31.2	100.0
Ecuador	114.4	93.2	+21.2	112.5
Guatemala	86.1	71.7	+14.4	84.4
Mexico	106.2	101.4	+4.8	106.1
Peru	101.7	80.5	+21.2	99.2
Trinidad & Tob.	93.8	33.3	+60.5	91.2

Note: Birth ratios are defined as $\frac{B5}{1/2\ (B4 + B6)} \times 100$

where B4, B5 and B6 are the number of births in the fourth, fifth and sixth calendar years before the survey.

[a]Ratio for living children less ratio for dead children
[b]Birth ratios, calculated here on the basis of true calendar periods, differ from those presented in Arnold, Table 3.3 (this volume), which are calculated for time periods relative to the cutoff dates for the collection of health-related data. These four surveys used cut-off dates that were not 1 January. Use of true calendar years is the appropriate method for assessing the impact of birth displacement on estimates of period mortality.

3.4 DATA ENTRY AND IMPUTATION PROGRAMS

While investigating the birth displacement problem, an error was discovered in the data processing programs used for the surveys in Mali, Thailand, Togo and Tunisia. The data entry program was designed to accept child health data only for births with a recorded date of birth the same as or later than the cut-off date specified in the questionnaire (e.g., January 1 of the fifth calendar year preceding the survey). Births for which the year of birth was not recorded did not meet this criterion; therefore, the health data recorded on the questionnaires were not entered. Later, at the imputation stage of data processing, the range of possible dates to which a birth (with missing year of birth) could be imputed, was limited to a period prior to the cut-off date if the health data were missing as well.[4] For some births, these data entry and imputation rules incorrectly precluded imputing a date of birth to the five-year period preceding the survey.

Relatively few births were affected by this programming error; however, births of children who died were disproportionately affected. This error raises the possibility of bias in the mortality estimates (i.e., negative bias for the five-year reference period immediately preceding the survey).

For the four surveys, birth dates were reimputed and the mortality rates recalculated, in order to determine the extent of bias in the original estimates. The results are shown in Table 3.3 for the five-year period preceding the survey. The revised estimates of infant mortality exceed the original estimates substantially in the case of Mali (9 percent), less so for Thailand (4 percent) and only slightly for Togo and Tunisia (1 percent). As expected, significant bias was found in the surveys where year of birth was missing for a high proportion of dead children (i.e., Mali and Thailand; see Table 3.1).

[4] An exception to this procedure occurred if the birth date of the preceding birth was after the cut-off date.

Table 3.3 Births in the five years preceding the survey, number of living and dead children, and infant mortality rates, based on standard recode data and reimputed data, Demographic and Health Surveys, 1986-1989

Country	Standard Recode Data	Reimputed Data	Percent Increase
Mali			
Total Births	3649	3754	2.9
Living Children	3062	3087	1.1
Dead Children	587	667	13.6
Infant Mortality Rate	108	118	9.3
Togo			
Total births	3251	3325	2.2
Living Children	2905	2947	1.4
Dead Children	346	378	9.2
Infant Mortality Rate	80.5	81.3	1.0
Tunisia			
Total Births	5084	5091	0.1
Living Children	4811	4812	0.0
Dead Children	273	279	2.2
Infant Mortality Rate	50.4	51.0	1.2
Thailand			
Total Births	3975	3983	0.2
Living Children	3828	3825	0.0
Dead Children	147	158	7.4
Infant Mortality Rate	35.6	37.0	3.9

4 Age at Death Data

In DHS surveys, age at death was collected in either days (for children dying within a month of birth), months (for children dying after one month but before their second birthday), or years (for children dying after their second birthday).[5] This information was recorded on the questionnaire by first circling the appropriate units code and then recording the reported number of days, months or years.

There are two variables in the standard recode data files which indicate age at death. The first variable preserves the information recorded on the questionnaire. It occupies a three-digit field: the first is the units indicator and the last two indicate the number of units. Dead children for which the data are missing or inconsistent receive a special code. The second variable recodes values of the first variable into completed months. Cases with special codes receive an imputed value using the hot deck procedure (i.e., the value of the last dead child of the same birth order).

Interviewers did not always follow the correct procedures for recording age at death. A response of "one year" was sometimes recorded rather than first probing for and then recording the age at death in months. In some surveys, values of "one year" were changed to 12 months during data entry, and that value appears for both age at death variables in the standard recode data file. In other surveys, values of age at death of "one year" were not systematically changed during data entry, and "one year" appears for the first variable in the recode data file and 12 months for the second variable.[6]

4.1 COMPLETENESS OF THE DATA

The age at death data are considered complete if both the units indicator and the number of units were recorded and the reported age at death was less than the interval between the child's date of birth and the date of the interview. Table 4.1 shows the percentage of deaths among children under five with incomplete data. Values are 1 percent or less in fifteen surveys. Incomplete data occurs for more than 2 percent of cases in only two surveys, Brazil (2.5 percent) and Mexico (5.1 percent). The high level of completeness in DHS surveys indicates that respondents readily reported this information or, that interviewers were able to estimate an age at death, after probing.

Table 4.1 Percentage of deaths among children under five with incomplete information on age at death by calendar period in which the death occurred, Demographic and Health Surveys, 1986-1989

	Calendar-Year Period Preceding the Survey					
Country	0-5	6-10	11-15	16-20	21-25	0-25
SUB-SAHARAN AFRICA						
Botswana	0.0	0.5	2.3	6.5	3.1	1.9
Burundi	0.7	1.0	0.6	3.4	0.8	1.1
Ghana	0.1	0.0	0.0	0.6	0.0	0.1
Liberia	1.3	1.3	0.6	1.0	0.3	1.0
Mali	0.3	0.4	0.0	0.4	0.0	0.3
Senegal	0.4	0.0	0.1	0.0	0.0	0.2
Togo	0.3	0.0	0.7	0.4	0.0	0.3
Uganda	0.8	0.2	0.2	0.0	0.0	0.4
Zimbabwe	1.4	0.9	0.0	0.6	1.0	0.8
NORTH AFRICA						
Morocco	0.4	0.5	0.2	0.0	0.2	0.3
Tunisia	2.7	0.5	1.2	1.8	2.5	1.7
ASIA						
Indonesia	0.3	0.1	0.0	0.1	0.0	0.1
Sri Lanka	0.5	0.0	0.0	0.6	0.0	0.2
Thailand	0.0	0.4	1.1	1.2	0.0	0.6
LATIN AMERICA/ CARIBBEAN						
Brazil	1.6	3.8	1.8	3.1	1.3	2.5
Colombia	0.6	0.5	1.0	0.6	0.9	0.7
Dominican Rep.	1.9	0.8	0.2	0.0	0.4	0.7
Ecuador	1.4	2.1	1.5	0.7	1.6	1.5
Guatemala	0.0	0.0	0.0	0.0	0.0	0.0
Mexico	2.4	5.3	4.9	6.7	6.3	5.1
Peru	0.0	0.0	0.0	0.0	0.0	0.0
Trinidad & Tob.	0.0	1.4	1.1	3.4	1.8	1.4

Note: The period in which a death occurred is derived from the date of birth and age at death. Where either piece of information is missing, imputed values are used.

It has generally been believed that survey respondents are less able to provide information about events which occurred in the distant past than about those which occurred

[5] The only exception was Mexico, where age at death (for deaths under one month) was not obtained in days.

[6] The extent to which age at death was recorded on the questionnaires as "one year" (and changed during editing to 12 months) cannot be determined from the recode data files for several countries. Confirmation that this occurred comes from reports by DHS staff. There are, however, eleven surveys for which the value "one year" was preserved in the standard recode data files for the first age at death variable (Burundi, Ghana, Guatemala, Liberia, Mali, Morocco, Senegal, Togo, Trinidad and Tobago, Tunisia, and Zimbabwe). In the following analysis of heaping of deaths at 12 months of age, no distinction is made between genuine digit preference (i.e., reports heaped at 12 months) and values of 12 months resulting from a conversion of "one year" responses.

Table 4.2 Total reported deaths among children under five and the number of deaths with incomplete age at death information by type of defect in information, Demographic and Health Surveys, 1986-1989

Country	Total Reported Deaths	Deaths Lacking Complete Age at Death Information	Defect in Age at Death Information		
			All Information Missing	Some Information Missing	Inconsistent Response
SUB-SAHARAN AFRICA					
Botswana	739	14	13	1	0
Burundi	2236	25	23	0	2
Ghana	2325	3	3	0	0
Liberia	4233	43	35	8	0
Mali	3301	9	5	4	0
Senegal	3335	5	0	5	0
Togo	1926	6	0	6	0
Uganda	2810	11	3	6	2
Zimbabwe	1128	9	7	2	0
NORTH AFRICA					
Morocco	3506	10	7	2	1
Tunisia	1664	28	18	10	0
ASIA					
Indonesia	4997	5	3	2	0
Sri Lanka	971	2	1	1	0
Thailand	1210	7	4	3	0
LATIN AMERICA/CARIBBEAN					
Brazil	1381	35	31	0	4
Colombia	832	6	6	0	0
Dominican Republic	2214	16	0	0	16
Ecuador	1328	20	14	0	6
Guatemala	1968	0	0	0	0
Mexico	1911	98	98	0	0
Peru	1739	0	0	0	0
Trinidad & Tobago	350	5	5	0	0

more recently. This premise would suggest that reporting is more complete for recent events than for those further in the past. Table 4.1 indicates, however, that such a pattern is not found in most DHS surveys.

Table 4.2 shows the distribution of deaths among children under five with incomplete data, by type of defect in the age at death information. In all but three surveys (Dominican Republic, Senegal, and Togo), the majority of deaths with incomplete data involve cases where both the units indicator and the number of units were missing. In the Dominican Republic, inconsistent responses account for all cases of incomplete data (16).[7]

4.2 ACCURACY OF THE DATA

Age-specific mortality estimates can be biased if misreporting of age at death results in a net transfer of deaths from one age interval to another. In retrospective surveys, heaping of reported age at death at 12 months is quite common. If heaping at 12 months is due to respondents rounding down the age of children who died after their first birthday, infant and child mortality estimates will be unaffected. If heaping is due to rounding up the age at death of late infant deaths, infant mortality will be underestimated and child mortality will be overestimated.

An index of heaping at 12 months was calculated as the number of deaths at 12 months of age divided by the average number of deaths at months 10, 11, 13 and 14. Under the assumption that the actual number of deaths changes linearly between 10 and 14 months, a value greater than 1.0 indicates heaping of deaths at 12 months. Table 4.3 shows the values of the heaping index. Values range from a low near 1.0 (Sri Lanka) to a high of 32 (Ghana). For the 22 surveys, the median value of the index is 11, which is identical to the value found in the assessment of the WFS data (Rutstein, 1985).

The idea that respondents are less able to provide precise information about events that occurred further in the past

[7] A reported age at death is inconsistent when it is greater than the interval between the reported birth date and the date of interview.

suggests that the heaping index should be greater for the more distant time periods. This pattern is evident in only three surveys (Indonesia, Peru, and Uganda), while the opposite is found in one survey (Colombia). No consistent pattern is evident in the remaining surveys.

Table 4.3 Index of heaping of deaths at twelve months of age by calendar period in which the death occurred, Demographic and Health Surveys, 1986-1989

	Calendar-Year Period Preceding the Survey					
Country	0-5	6-10	11-15	16-20	21-25	0-25
SUB-SAHARAN AFRICA						
Botswana	12.6	20.7	10.4	14.0	a	14.4
Burundi	29.6	18.5	47.8	34.2	13.8	26.2
Ghana	39.1	30.0	41.3	21.7	27.0	32.4
Liberia	17.5	13.5	18.4	17.0	10.1	15.5
Mali	21.8	21.5	26.8	35.0	28.9	24.8
Senegal	15.9	21.0	25.2	26.7	23.5	21.3
Togo	18.9	13.5	49.0	26.4	12.6	19.9
Uganda	9.9	20.8	19.4	22.3	25.8	15.8
Zimbabwe	9.2	10.6	13.3	10.0	17.0	11.3
NORTH AFRICA						
Morocco	11.6	10.3	10.8	11.7	33.7	12.4
Tunisia	6.8	33.6	78.0	34.4	29.3	26.4
ASIA						
Indonesia	3.4	8.2	9.6	10.9	7.3	7.5
Sri Lanka	0.2	1.1	1.7	0.0	0.5	0.7
Thailand	15.3	9.3	14.3	14.1	13.3	12.8
LATIN AMERICA/CARIBBEAN						
Brazil	3.4	1.6	5.2	5.0	6.1	3.4
Colombia	11.4	7.8	6.9	3.7	2.7	6.0
Dominican Rep.	6.6	5.0	5.8	4.6	5.0	5.3
Ecuador	9.8	16.0	8.4	11.3	8.0	10.5
Guatemala	7.4	11.5	12.0	10.3	7.4	9.7
Mexico	2.8	5.5	4.4	3.1	6.0	4.1
Peru	9.1	6.7	7.3	16.4	30.7	9.9
Trinidad & Tob.	4.0	0.0	8.0	2.0	12.0	4.0

Note: The index of heaping was calculated as:

$$\frac{(4xD12)}{(D10+D11+D13+D14)}$$ where D12 includes all deaths reported at 12 months and 1 year.

[a]Five deaths reported at 12 months; no deaths in adjacent months.

4.3 IMPACT ON MORTALITY ESTIMATES

When age at death data are heaped at 12 months, there is no certain procedure to determine the proportion that are infant deaths and the proportion that are deaths at age one (i.e., age 12-23 months). Some analysts argue that heaping reflects, in large part, a rounding down of deaths at age one and they make no adjustments for it (Rutstein, 1984). Other analysts assume that a significant proportion of deaths are drawn from infancy and they reassign as many as half of all deaths recorded at 12 months to infancy (Goldman et. al., 1979; Thapa and Retherford, 1982). The mortality rates presented in the DHS survey reports are not adjusted for heaping at 12 months. However, in many datasets it is clear that there is a shortfall of deaths in late infancy, indicating that at least part of the heaping at 12 months reflects rounding up the reported age at death (see Figure 4.1).

A measure of the impact of heaping on mortality estimates can be obtained by reassigning a proportion of the deaths at 12 months to infancy and recalculating mortality rates. The results from a model which redistributes 25 percent of the excess deaths at 12 months to infancy are shown in Table 4.4.[8] In terms of infant mortality, the adjustment increases the estimated rates less for surveys in North Africa, Asia, and Latin America than for surveys in sub-Saharan Africa, where heaping is greater. The mean increase in estimates of the infant mortality rate (IMR) for surveys in sub-Saharan Africa is 5 percent, compared with 2 percent for surveys in the other regions. In all but three surveys the adjustment procedure produced smaller percent increases in infant mortality than percent decreases in child mortality. In only three African countries, which have very high childhood mortality (Burundi, Mali, and Senegal), was the percent increase in infant mortality less than the decrease in child mortality.

[8]The model calculates excess deaths as the difference between the number of deaths at 12 months and the average number at months 10, 11, 13, and 14. This procedure is conservative in the sense that it tends to overestimate the number of excess deaths at 12 months of age.

Figure 4.1 Age at death for children who died at less than two years of age, selected Demographic and Health Surveys, 1987

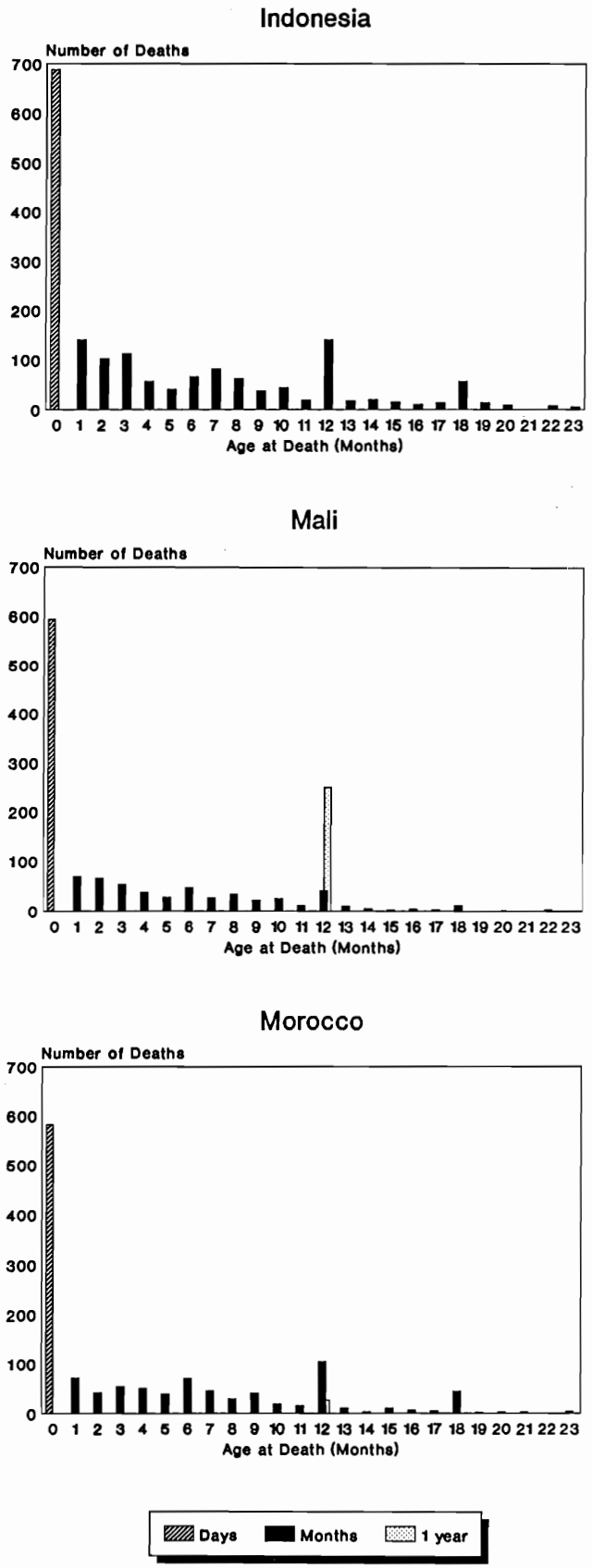


Table 4.4 Estimates of infant and child mortality for the ten-year period preceding the survey, adjusted for heaping of deaths at twelve months of age, Demographic and Health Surveys, 1986-1989

Country	Infant Mortality ($_1q_0$)			Child Mortality ($_4q_1$)		
	Unadjusted Rate	Adjusted Rate	Percent Increase	Unadjusted Rate	Adjusted Rate	Percent Decrease
SUB-SAHARAN AFRICA						
Botswana	39.5	41.2	4.3	17.0	15.3	10.3
Burundi	87.1	92.1	5.8	107.8	102.7	4.8
Ghana	81.3	86.0	5.8	78.9	74.0	6.2
Liberia	151.6	158.1	4.3	92.2	84.9	7.9
Mali	132.5	142.7	7.7	169.7	159.6	6.0
Senegal	90.1	96.5	7.1	131.5	25.1	4.9
Togo	83.2	87.5	5.1	82.3	77.9	5.4
Uganda	105.9	110.5	4.4	90.6	85.6	5.6
Zimbabwe	57.4	59.1	3.0	31.1	29.3	5.8
NORTH AFRICA						
Morocco	82.4	84.8	2.9	38.7	36.2	6.5
Tunisia	56.8	58.3	2.1	18.8	17.2	8.5
ASIA						
Indonesia	75.2	77.0	2.4	39.1	37.6	4.0
Sri Lanka	32.5	32.5	0.0	10.2	10.2	0.0
Thailand	38.3	38.8	1.3	11.0	10.4	4.5
LATIN AMERICA/CARIBBEAN						
Brazil	86.2	86.5	0.4	13.6	13.2	2.5
Colombia	39.3	40.4	2.6	12.8	11.7	8.6
Dominican Republic	70.4	71.4	1.5	23.9	22.7	4.8
Ecuador	65.2	66.9	2.6	26.4	24.6	7.0
Guatemala	79.5	82.6	3.9	45.9	42.6	7.1
Mexico	56.2	56.9	1.2	15.5	14.9	4.2
Peru	79.2	81.8	3.2	38.8	36.1	6.9
Trinidad & Tobago	31.1	31.2	0.3	3.4	3.4	0.0

Note: Estimates are adjusted by reassigning 25 percent of the "excess" deaths reported at 12 months back to the 6-11 age segment. (See text for a definition of "excess" deaths.)

5 Completeness of Event Reporting

Potentially the most serious source of error in the birth history data is the omission of children who have died. In this section, event omission is investigated by checking the internal consistency of the reported data and by comparing DHS estimates with those from WFS surveys.

5.1 INTERNAL CONSISTENCY

The rationale for internal consistency tests is that mortality rates typically follow well-established patterns when examined according to certain characteristics (age, sex, birth order, etc.) and that departures from the expected patterns are indicative of defective data. However, it should be recognized at the outset that internal consistency tests cannot detect underreporting which is moderate in degree or which is non-selective across values for the characteristic being investigated. Therefore, while substantial departures from expected patterns can signal data errors, the absence of such findings does not confirm that reporting of events is complete.

Age Pattern of Mortality

Mortality rates typically decline sharply in the first few days and weeks of life and with few exceptions continue to decline, although less sharply, through late infancy and early childhood. It is also known that the steepness of the decline in mortality risk with age is related to the level of mortality; typically, at lower mortality levels deaths are more concentrated at younger ages, so that the pattern of decline with age is more pronounced than at higher mortality levels.

Table 5.1 Infant mortality rates and ratio of neonatal to infant mortality by calendar year period preceding the survey, Demographic and Health Surveys, 1986-1989

	Infant Mortality Rate by Calendar-Year Period Preceding the Survey			Ratio of Neonatal Mortality to Infant Mortality by Calendar-Year Period Preceding the Survey		
Country	0-5	6-10	11-15	0-5	6-10	11-15
SUB-SAHARAN AFRICA						
Liberia	143.7	161.4	192.2	.47	.45	.43
Mali	108.4	157.1	166.9	.49	.55	.55
Uganda	100.7	112.9	86.9	.45	.48	.45
Senegal	85.8	95.1	117.6	.52	.59	.48
Togo	80.5	96.7	107.4	.54	.53	.50
Ghana	77.2	86.4	99.6	.57	.59	.63
Burundi	74.5	102.9	100.2	.47	.42	.43
Zimbabwe	52.7	63.7	53.6	.52	.55	.59
Botswana	37.4	42.1	57.3	.58	.57	.45
NORTH AFRICA						
Morocco	73.3	92.4	104.4	.56	.53	.47
Tunisia	50.4	64.4	70.0	.55	.49	.46
ASIA						
Indonesia	70.2	80.5	80.7	.58	.57	.51
Thailand	35.6	40.9	54.6	.56	.68	.69
Sri Lanka	25.3	39.4	39.3	.64	.63	.58
LATIN AMERICA/CARIBBEAN						
Peru	75.8	83.3	91.3	.44	.43	.41
Brazil	74.7	100.2	99.7	.44	.43	.49
Guatemala	73.3	87.6	93.7	.44	.47	.51
Dominican Republic	67.2	74.3	79.3	.58	.57	.51
Ecuador	58.0	72.5	95.2	.61	.55	.53
Mexico	46.9	66.6	70.2	.56	.50	.56
Colombia	33.4	46.6	60.7	.58	.48	.52
Trinidad & Tobago	26.2	37.2	46.7	.81	.67	.61

Note: Within region, countries are listed in order of decreasing infant mortality.

It is often hypothesized that children who die soon after birth are those most likely to be unreported in a retrospective survey. If such selective underreporting is severe, an abnormally low ratio of neonatal to infant mortality would be observed. Table 5.1 shows these ratios for three five-year periods before each survey. The values of the ratio range from 0.41 (Peru) to 0.81 (Trinidad and Tobago). As expected, lower levels of mortality are associated with higher values of this ratio. For a given country, there is a tendency for the ratio to be higher for more recent time periods, particularly if mortality has declined over time. Exceptions to this pattern are Guatemala, Mali, Thailand, and Zimbabwe.

The possibility of underreporting of early infant deaths can be further investigated by examining the distribution of deaths within the neonatal period. Table 5.2 presents statistics on the ratio of deaths under seven days (early neonatal period) to all neonatal deaths. In general, the ratios do not indicate underreporting of births of children who died in early infancy. The ratios are 0.52 or higher and are frequently in excess of 0.70.[9]

The analyses of mortality patterns by age has not detected evidence of substantial underreporting of deaths in the DHS surveys, with the possible exception of Thailand. The Thailand data indicate a decline in neonatal mortality of nearly 50 percent in the period between 1972-76 and 1982-87; at the same time, postneonatal mortality remained virtually unchanged.

Sex Differentials

There are reasons to expect variations in mortality risk between male and female children. Biological factors predispose boys to higher risk, especially during infancy. Behavioral factors may operate in the opposite direction in

[9] There are no model mortality patterns for the neonatal period. However, one review of data from several developing countries concluded that, at levels of neonatal mortality of 20 per 1000 or higher, approximately 70 percent of neonatal deaths occur within the first six days of life (Boerma, 1988).

Table 5.2 Neonatal mortality rates and the ratio of deaths under seven days to all neonatal deaths by calendar year period preceding the survey, Demographic and Health Surveys, 1986-1989

	Neonatal Mortality Rate by Calendar-Year Period Preceding the Survey			Ratio of Deaths Under 7 Days to All Neonatal Deaths by Calendar-Year Period Preceding the Survey		
Country	0-5	6-10	11-15	0-5	6-10	11-15
SUB-SAHARAN AFRICA						
Liberia	67.8	73.4	82.8	.69	.66	.61
Mali	53.6	87.0	91.1	.68	.68	.65
Uganda	45.5	54.6	43.8	.67	.74	.74
Senegal	44.2	52.0	55.9	.60	.55	.57
Togo	43.4	46.0	53.6	.74	.73	.69
Ghana	43.8	51.7	62.4	.77	.69	.75
Burundi	34.9	43.4	43.2	.68	.67	.59
Zimbabwe	26.6	35.3	32.9	.77	.57	.62
Botswana	21.6	23.9	26.0	.81	.76	.76
NORTH AFRICA						
Morocco	41.1	48.8	48.8	.52	.52	.49
Tunisia	27.8	27.8	32.5	.69	.72	.65
ASIA						
Indonesia	30.0	41.9	37.6	.52	.48	.52
Thailand	20.1	27.7	37.9	.74	.65	.62
Sri Lanka	16.3	24.7	22.9	.72	.76	.73
LATIN AMERICA/CARIBBEAN						
Peru	33.6	35.9	37.8	.58	.60	.51
Brazil	33.0	43.4	49.4	.61	.56	.58
Guatemala	32.9	41.4	47.6	.70	.78	.75
Dominican Republic	38.7	42.5	40.2	.70	.77	.71
Ecuador	35.4	39.8	50.4	.65	.64	.55
Mexico	26.0	33.4	39.0	NA	NA	NA
Colombia	19.4	22.3	31.8	.79	.73	.73
Trinidad & Tobago	21.3	25.0	28.6	.76	.86	.70

Note: Within region, countries are listed in order of decreasing infant mortality.

societies where there is a strong preference for male children and child care practices differ by sex.

The IMR sex ratio (male IMR/female IMR) of the Regional Model Life Tables varies from 1.36 to 1.16 for levels of infant mortality from 25 to 200 per 1000 (Coale and Demeny, 1966). These values were used to define a plausible range for the IMR sex ratio in the assessment of the WFS mortality data (Rutstein, 1985) and are used for that purpose here.

If births of female children who have died are omitted more frequently than births of male children who have died (as is often hypothesized), then the IMR sex ratio will be positively biased. Less likely, if in a particular population there is a reluctance to disclose deaths of more culturally-valued male children, the ratio will be negatively biased. Table 5.3 shows infant mortality rates by sex, and the IMR sex ratio for the ten-year period preceding the survey. For most surveys, the value of the IMR sex ratio is within the plausible range. However, the values exceed the range in three surveys and fall below it in three other surveys.

Table 5.3 Infant mortality rates by sex of child and sex risk ratio (M/F) for the ten-year period preceding the survey, Demographic and Health Surveys, 1986-1989

Country	Infant Mortality Rate		Sex Risk Ratio (M/F)
	Males	Females	
SUB-SAHARAN AFRICA			
Botswana	47.7	31.5	1.51
Burundi	98.2	75.7	1.30
Ghana	88.8	73.5	1.21
Liberia	167.2	135.4	1.23
Mali	138.2	126.5	1.09
Senegal	97.3	82.5	1.18
Togo	87.7	78.5	1.12
Uganda	111.3	100.5	1.11
Zimbabwe	64.4	49.4	1.30
NORTH AFRICA			
Morocco	83.4	81.1	1.03
Tunisia	58.0	55.5	1.05
ASIA			
Indonesia	84.2	65.6	1.28
Sri Lanka	39.6	24.8	1.60
Thailand	45.1	31.1	1.45
LATIN AMERICA/CARIBBEAN			
Brazil	98.5	73.4	1.34
Colombia	40.7	37.9	1.07
Dominican Republic	78.7	61.8	1.27
Ecuador	70.1	60.0	1.17
Guatemala	90.6	68.0	1.33
Mexico	60.1	52.4	1.15
Peru	83.7	74.5	1.12
Trinidad & Tobago	28.8	33.5	0.86

For Botswana, Sri Lanka, and Thailand the IMR sex ratios exceed 1.40, suggesting that there was underreporting of female infants who had died. If omission did take place, it is likely to have occurred for deaths in the neonatal period, which would be indicated by particularly high ratios of male to female neonatal mortality rates. In the case of Botswana and Sri Lanka, very high values were found for the neonatal period (1.94 and 1.76, respectively) and more moderate values for the postneonatal period (1.09 and 1.45, respectively). In the case of Thailand, the mortality sex ratios for the two periods of infancy were identical (1.45).

For Tunisia, Morocco, and Trinidad and Tobago the IMR sex ratios were 1.05 or lower, indicating possible omission of male deaths. The IMR sex ratios for the WFS surveys in Tunisia and Morocco were similarly low (1.02 and 1.06, respectively) while the ratio for Trinidad and Tobago (1.20) was in the acceptable range (Rutstein, 1985). Examination of the sex differentials for subintervals of infancy found that the low IMR sex ratios for Tunisia and Morocco are the result of high female (relative to male) mortality for the postneonatal period (which is consistent with child care practices favoring males), and that, for the neonatal period, male mortality exceeds female mortality, as expected. However, the case of Trinidad and Tobago is not as readily explained, since female mortality exceeds male mortality in both the neonatal and postneonatal periods.

Birth Order and Maternal Age

These two variables are highly correlated and their relation to childhood mortality is expected to show similar patterns. It is usually observed that as both birth order and maternal age increase the rate of mortality first falls and then rises. Selective omission of deaths_for example, by first-time mothers or very young mothers_could disturb this pattern. Underreporting of first births that have died is thought to be particularly likely, especially if the birth occurred before entry into a stable union.

Infant mortality rates and relative risks by birth order and maternal age are shown in Tables 5.4 and 5.5, respectively. In most surveys there are no marked departures from the expected pattern; however, for Botswana, Brazil, Colombia, Mexico, and Thailand, the risk for first births is (contrary to expectation) well below that for second and third order births. In all of these countries except Botswana, young maternal age is associated with higher mortality risk. The relatively low risk of mortality for first births is probably not indicative of underreporting; rather, it is possible that later entry into marriage could have resulted in a significant proportion of first births occurring at more favorable ages.

In the case of Botswana, relative risk is low both for first births and for births occurring before age 20. Furthermore,

Table 5.4 Infant mortality rates by birth order of child, for the ten-year calendar period preceding the survey, Demographic and Health Surveys, 1986-1989

Country	Infant Mortality Rate by Birth Order of Child				Relative Risk of Death by Birth Order of Child[a]			
	1	2-3	4-6	7+	1	2-3	4-6	7+
SUB-SAHARAN AFRICA								
Botswana	32.4	44.6	37.5	44.3	0.73	1.00	0.84	0.99
Burundi	98.5	84.6	80.9	93.0	1.16	1.00	0.96	1.10
Ghana	86.3	67.9	82.6	101.8	1.27	1.00	1.22	1.50
Liberia	154.4	146.5	146.4	169.4	1.05	1.00	1.00	1.16
Mali	170.3	119.4	113.2	150.7	1.43	1.00	0.95	1.26
Senegal	111.9	84.9	81.4	91.1	1.32	1.00	0.96	1.07
Togo	79.5	83.6	79.7	93.6	0.95	1.00	0.95	1.12
Uganda	120.0	102.7	104.7	99.1	1.17	1.00	1.02	0.96
Zimbabwe	61.8	53.5	47.6	76.2	1.16	1.00	0.89	1.42
NORTH AFRICA								
Morocco	90.7	80.1	78.4	84.8	1.13	1.00	0.98	1.06
Tunisia	60.7	44.2	60.1	76.0	1.37	1.00	1.36	1.72
ASIA								
Indonesia	78.1	70.3	70.5	94.0	1.11	1.00	1.00	1.34
Sri Lanka	31.1	29.6	37.5	46.5	1.05	1.00	1.27	1.57
Thailand	29.8	36.1	48.0	72.8	0.83	1.00	1.32	2.01
LATIN AMERICA/CARIBBEAN								
Brazil	62.0	73.0	100.4	159.8	0.85	1.00	1.38	2.19
Colombia	30.4	41.0	42.4	51.1	0.74	1.00	1.03	1.25
Dominican Republic	69.5	64.1	68.2	91.2	1.08	1.00	1.06	1.42
Ecuador	58.6	52.3	68.8	103.9	1.12	1.00	1.32	1.99
Guatemala	84.3	70.6	78.1	99.8	1.19	1.00	1.11	1.41
Mexico	42.0	50.3	60.0	87.9	0.83	1.00	1.19	1.75
Peru	61.6	62.9	89.3	115.0	0.98	1.00	1.42	1.83
Trinidad & Tobago	28.9	25.6	33.0	72.6	1.13	1.00	1.29	2.84

[a]Reference category is birth orders 2-3.

Table 5.5 Infant mortality rates by age of mother at birth of child, for the ten-year calendar period preceding the survey, Demographic and Health Surveys, 1986-1989

Country	Infant Mortality Rate by Age of Mother at Birth of Child				Relative Risk of Death by Age of Mother at Birth of Child[a]			
	<20	20-29	30-39	40+	<20	20-29	30-39	40+
SUB-SAHARAN AFRICA								
Botswana	34.9	41.8	38.8	(35.8)	0.83	1.00	0.93	0.86
Burundi	137.9	86.5	75.6	(74.4)	1.59	1.00	0.87	0.86
Ghana	97.0	73.1	82.8	(118.6)	1.32	1.00	1.13	1.62
Liberia	175.9	145.7	135.6	(170.8)	1.21	1.00	0.93	1.17
Mali	177.8	116.6	122.1	(178.2)	1.52	1.00	1.05	1.53
Senegal	117.7	82.4	82.4	(98.1)	1.43	1.00	1.00	1.19
Toto	90.3	79.1	87.9	(76.3)	1.14	1.00	1.11	0.96
Uganda	119.4	104.2	93.0	(129.7)	1.15	1.00	0.89	1.24
Zimbabwe	76.5	47.5	62.3	(67.1)	1.61	1.00	1.31	1.41
NORTH AFRICA								
Morocco	122.5	79.1	76.2	65.7	1.55	1.00	0.96	0.83
Tunisia	68.5	55.6	51.9	(109.5)	1.23	1.00	0.93	1.97
ASIA								
Indonesia	99.2	68.1	74.2	71.1	1.45	1.00	1.09	1.04
Sri Lanka	34.6	33.4	28.8	(51.9)	1.03	1.00	0.86	1.55
Thailand	39.5	33.4	47.4	(68.8)	1.18	1.00	1.42	2.06
LATIN AMERICA/CARIBBEAN								
Brazil	103.0	81.5	86.0	(142.8)	1.26	1.00	1.05	1.75
Colombia	46.1	36.7	40.1	(45.0)	1.26	1.00	1.09	1.23
Dominican Republic Republic	84.6	60.5	79.1	(104.6)	1.40	1.00	1.31	1.73
Ecuador	61.9	63.4	62.9	(140.8)	0.98	1.00	0.99	2.22
Guatemala	98.9	72.5	77.7	(165.6)	1.36	1.00	1.07	2.28
Mexico	63.1	53.0	57.1	(73.9)	1.19	1.00	1.08	1.39
Peru	100.2	68.1	85.9	(103.6)	1.47	1.00	1.26	1.52
Trinidad & Tobago	42.9	28.4	25.5	(91.6)	1.51	1.00	0.89	3.22

Note: Figures in parentheses are based on fewer than 500 observations.

[a]Reference category is births to women 20-29.

the overall pattern of mortality by birth order and maternal age is atypical in Botswana, since relative risk is highest for second and third order births and for women 20 to 29. One interpretation of these findings is that dead children who were born when the respondents were young, first-time mothers are underreported.

Preceding Birth Interval

Infant and child mortality risk has consistently been shown to vary with the length of the preceding birth interval. That is, births which follow a short interval (less than 24 months) are at higher risk than those which follow a longer interval (24-47 and 48+ months). Table 5.6 shows infant mortality rates by birth interval. In all surveys, short intervals are associated with increased risk of infant mortality relative to the 48+ month reference category. Relative risk ranges from over 3.0 (Burundi, Mali, and Peru) to about 1.3 (Sri Lanka and Trinidad and Tobago). The low values of relative risk for Sri Lanka and Trinidad and Tobago are not necessarily indicative of data problems, since the risk for the short interval (<24 months) is quite high relative to the long interval (24-47 months).

Maternal Education

It is well known that infant and child mortality are inversely related to the level of maternal education. The strength of the relationship, however, varies between countries. In Table 5.7, infant mortality rates and relative risk are shown by level of maternal education. The reader should bear in mind that the three education categories (none, primary, and secondary or higher) are not absolutely comparable across the 22 surveys.

The effect of maternal education on infant mortality varies widely among the countries surveyed. The expected pattern of declining mortality with increasing education is evident

Table 5.6 Infant mortality rates by length of retrospective birth interval for the ten-year calendar period preceding the survey, Demographic and Health Surveys, 1986-1989

Country	Infant Mortality Rate by Length of Birth Interval (Months)			Relative Risk of Death by Length of Birth Interval (Months)[a]		
	<24	24-47	48+	<24	24-47	48+
SUB-SAHARAN AFRICA						
Botswana	69.7	37.5	33.5	2.08	1.12	1.00
Burundi	133.2	72.3	41.1	3.24	1.78	1.00
Ghana	121.2	72.4	66.6	1.82	1.09	1.00
Liberia	214.4	132.9	87.8	2.43	1.51	1.00
Mali	206.8	114.5	53.9	3.84	2.12	1.00
Senegal	120.0	75.6	69.2	1.73	1.09	1.00
Togo	126.1	77.3	71.7	1.76	1.08	1.00
Uganda	143.0	87.1	79.3	1.80	1.10	1.00
Zimbabwe	81.3	50.4	48.8	1.67	1.05	1.00
NORTH AFRICA						
Morocco	120.8	63.4	49.7	2.43	1.28	1.00
Tunisia	89.4	34.5	36.4	2.46	0.95	1.00
ASIA						
Indonesia	110.7	63.5	51.8	2.14	1.23	1.00
Sri Lanka	45.5	25.1	35.3	1.29	0.71	1.00
Thailand	59.3	38.3	34.1	1.74	1.12	1.00
LATIN AMERICA/CARIBBEAN						
Brazil	143.4	61.4	55.5	2.58	1.11	1.00
Colombia	53.0	41.6	29.5	1.80	1.41	1.00
Dominican Republic	88.4	59.2	59.9	1.48	0.99	1.00
Ecuador	95.0	52.6	47.0	2.02	1.12	1.00
Guatemala	123.3	53.4	58.9	2.09	0.91	1.00
Mexico	83.9	48.3	46.2	1.82	1.05	1.00
Peru	124.2	66.7	37.2	3.34	1.79	1.00
Trinidad & Tobago	46.0	16.8	32.0	1.44	0.53	1.00

[a]Reference category is births with preceding birth interval 24-47 months in length.

Table 5.7 Infant mortality rates by mother's education for the ten-year calendar period preceding the survey, Demographic and Health Surveys, 1986-1989

Country	Infant Mortality Rate by Mother's Education			Relative Risk of Death by Mother's Education[a]		
	None	Primary	Secondary or higher	None	Primary	Secondary or higher
SUB-SAHARAN AFRICA						
Botswana	46.5	35.4	34.7	1.34	1.02	1.00
Burundi	89.4	81.9	(32.3)	2.77	2.54	1.00
Ghana	87.7	75.0	(79.1)	1.11	0.95	1.00
Liberia	162.0	144.8	111.2	1.46	1.30	1.00
Mali	140.6	103.9	(84.9)	1.66	1.22	1.00
Senegal	95.5	65.2	(49.9)	1.91	1.31	1.00
Togo	87.3	78.4	(54.5)	1.60	1.44	1.00
Uganda	114.4	101.5	83.5	1.37	1.22	1.00
Zimbabwe	78.0	54.2	39.9	1.95	1.36	1.00
NORTH AFRICA						
Morocco	86.1	54.0	65.3	1.32	0.83	1.00
Tunisia	63.5	50.1	35.1	1.81	1.43	1.00
ASIA						
Indonesia	98.8	76.6	37.4	2.64	2.05	1.00
Sri Lanka	52.4	34.0	27.9	1.88	1.22	1.00
Thailand	53.7	38.7	19.0	2.83	2.04	1.00
LATIN AMERICA/CARIBBEAN						
Brazil	115.3	91.2	24.9	4.63	3.66	1.00
Colombia	54.8	42.0	28.8	1.90	1.46	1.00
Dominican Republic	98.9	73.5	47.3	2.09	1.55	1.00
Ecuador	106.2	68.3	39.5	2.69	1.73	1.00
Guatemala	82.2	81.9	40.5	2.03	2.02	1.00
Mexico	82.7	57.8	27.4	3.02	2.11	1.00
Peru	118.9	87.5	41.8	2.84	2.09	1.00
Trinidad & Tobago	(65.6)	24.7	39.6	1.66	0.63	1.00

Note: Figures in parentheses are based on fewer than 500 observations.

[a]Reference category is births to women educated to secondary level or higher.

in all countries except Morocco and Trinidad and Tobago. These two countries exhibit a pattern that may signal selective underreporting of deaths by women with less than a secondary education.

5.2 COMPARISON OF DHS DATA WITH OTHER SOURCES

The WFS surveys, conducted in the 1970s and early 1980s, are the best single source of national-level mortality data against which the DHS data can be evaluated. Of the 22 DHS surveys considered in this report, 13 are in countries where WFS surveys were carried out seven to thirteen years earlier.

In this analysis, the DHS and WFS mortality estimates—which are based on direct estimation techniques—are compared for a fixed time period. The WFS estimates are for the five-year period immediately preceding the survey; the DHS estimates are for that same time period (i.e., ten to fifteen years preceding the survey).[10] Thus, the DHS mortality estimates are based on events which require longer respondent recall (five times longer, on average) and which might be expected to be reported less completely.

[10] The WFS estimates had to be adjusted to compensate for the possible effect of truncation bias in the DHS estimates. Truncation bias arises from the unavailability of births to older women for time periods prior to the survey date. For instance, since the maximum age of survey respondents is 49, the DHS birth history data for a five-year period which is ten to fifteen years before the survey includes no births to women older than 39. If births to women 40 years and older have relatively high mortality, then the DHS estimate for that period will be an underestimate. Truncation bias was compensated for by estimating rates for women 15-34 in the WFS surveys. In the case of Indonesia, an additional step was required in order to establish geographical comparability for the WFS and DHS estimates: the DHS estimates were restricted to Java-Bali, which was the only area covered by the WFS survey.

Table 5.8 shows under-five mortality rates for the 13 countries where both WFS and DHS surveys were conducted. In all three Asian countries, the WFS mortality rates are more than 10 percent higher than the DHS rates: Indonesia (11 percent), Thailand (15 percent), and Sri Lanka (31 percent). In the six Latin American countries, the estimates differ substantially only for Colombia, where the WFS rate is 19 percent higher than the DHS rate. In the four African countries, there is no substantial difference between the estimates, except for Ghana, where the DHS estimate is 38 percent higher than the WFS estimate.

Table 5.8 Estimates of under-five mortality for a fixed five-year reference period, WFS and DHS Surveys

Country	Midpoint of Reference Period	Under-Five Mortality ($_5q_0$) WFS	Under-Five Mortality ($_5q_0$) DHS	Percent Difference[a]
SUB-SAHARAN AFRICA				
Ghana	1977.2	120.1	166.1	-38
Senegal	1976.0	261.4	258.9	+ 1
NORTH AFRICA				
Morocco	1977.7	143.6	154.9	- 8
Tunisia	1976.1	103.1	100.8	+ 2
ASIA				
Indonesia[b]	1973.8	157.6	139.7	+11
Sri Lanka	1973.3	84.6	58.6	+31
Thailand	1972.9	87.9	74.4	+15
LATIN AMERICA/CARIBBEAN				
Colombia	1974.0	104.8	84.4	+19
Dominican Republic	1972.9	126.8	117.0	+ 8
Ecuador	1977.3	115.5	110.1	+ 5
Mexico	1974.4	94.3	93.1	+ 1
Peru	1975.5	144.2	131.5	+ 9
Trinidad & Tobago	1974.9	50.3	52.8	- 5

Note: Estimates are for women 15-39.

[a] $\frac{\text{WFS-DHS}}{\text{WFS}}$

[b] Estimates do not include areas outside Java and Bali.

The shortfall (by more than 10 percent) in the DHS mortality estimates for four countries suggests underreporting of dead children for the period 10 to 15 years preceding the DHS survey. However, the findings do not necessarily imply underreporting of events for more recent time periods. Similarly, the absence of a shortfall in the other DHS surveys does not mean that reporting was necessarily complete in those surveys. It is possible that there was underreporting of events and underestimation of rates for the period immediately preceding the WFS survey. In the case of Ghana, the most plausible explanation is that there was underreporting of events in the WFS survey.

The findings do suggest that in analyzing trends in mortality, a cautious approach is warranted when data are available from only one survey. Actual mortality declines may be masked or underestimated if greater underreporting of events occurs for periods in the more distant past.

Table 5.9 shows DHS and WFS infant and child mortality estimates for the same five-year reference period. For the DHS surveys in which under-five mortality estimates are significantly lower than WFS estimates, the percent shortfall is greater for child mortality than for infant mortality, particularly in the case of Thailand. This pattern is contrary to expectation.

Table 5.9 Estimates of infant and child mortality for a fixed five-year reference period, WFS and DHS surveys

Country	Midpoint of Reference Period	Infant Mortality Rate ($_1q_0$)			Child Mortality Rate ($_4q_1$)		
		WFS	DHS	Percent Difference[a]	WFS	DHS	Percent Difference[a]
SUB-SAHARAN AFRICA							
Ghana	1977.2	69.7	88.1	-26	54.1	85.5	-58
Senegal	1976.0	112.2	109.9	+ 2	167.1	167.4	0
NORTH AFRICA							
Morocco	1977.7	91.1	105.9	-16	57.8	54.9	+ 5
Tunisia	1976.1	75.4	70.3	+ 7	30.0	32.8	- 1
ASIA							
Indonesia[b]	1973.9	93.1	83.2	+11	71.2	61.7	+13
Sri Lanka	1973.3	58.0	40.6	+30	27.9	18.8	+33
Thailand	1972.9	62.7	60.9	+ 3	26.8	14.4	+46
LATIN AMERICA/CARIBBEAN							
Colombia	1974.0	68.7	57.1	+17	39.3	29.0	+26
Dominican Republic	1972.9	86.4	76.2	+12	44.4	44.1	+ 1
Ecuador	1977.3	74.9	76.1	- 2	43.8	36.8	+16
Mexico	1974.4	70.2	71.0	- 1	25.8	23.8	+ 8
Peru	1975.5	92.5	87.6	+ 5	56.9	48.1	+15
Trinidad & Tobago	1974.9	42.8	46.2	- 8	7.9	7.0	+11

Note: Estimates are adjusted for births to women 15-39.

[a] $\frac{\text{WFS-DHS}}{\text{WFS}}$

[b] Estimates do not include areas outside Java and Bali.

6 Summary and Recommendations

This assessment of DHS mortality data first considered the quality of the reported birth history data and then addressed the more difficult issue of underreporting of events. The analysis of the reported data focused on the completeness and accuracy of the data on date of birth and age at death. In the case of underreporting of events, the analysis utilized internal consistency tests and comparisons of DHS data with data from other sources. Overall, the assessment found that the quality of the DHS data is good, although some data problems were detected. It is probable that modification of the procedures used in DHS surveys would reduce the effects of these data problems.

6.1 QUALITY OF THE REPORTED DATA

Analysis of the data on date of birth found that year of birth was missing much more frequently for dead children than for living children. For the fifteen-year period preceding a survey, the percentage of dead children requiring imputation of year of birth was 2 percent or more in half of the DHS surveys and over 25 percent in two surveys (Mali and Thailand). Regarding the displacement of births from the fifth to the sixth calendar year prior to the survey, in twelve surveys birth displacement was greater for dead children than for living children—a pattern which tends to underestimate mortality for the five-year reference period immediately preceding a survey. However, in most surveys, the impact on estimates of infant mortality was not great. A simulation estimated the potential impact to be between one and four percent.

Analysis of the data on age at death, found that heaping on 12 months of age was about as prevalent in DHS surveys as it was in previous WFS surveys. Overall, about ten times as many deaths were reported at 12 months of age as at age 10, 11, 13 or 14 months. Heaping was particularly pronounced in the surveys in sub-Saharan Africa. The impact on mortality estimates was determined by distributing 25 percent of the excess deaths at 12 months to infancy and recalculating the rates. Increases in the estimates of infant mortality averaged 2 percent for surveys in North Africa, Asia, and Latin America and about 5 percent for surveys in sub-Saharan Africa.

Mortality rates by demographic characteristics of the mother and child were analyzed in an effort to identify patterned differentials. Most surveys showed the expected variations; however, some raised questions about possible data problems.

- The mortality data from Botswana are suspect on two tests of internal consistency, indicating underreporting of births of children who died in the neonatal period, particularly female births and births to young, first-time mothers.
- Female neonatal deaths may have been underreported in Sri Lanka.
- There is a deficit of male neonatal deaths in the ten-year period preceding the Trinidad and Tobago survey.

In 13 countries, it was possible to compare DHS and WFS mortality estimates. For Ghana, the DHS estimate of under-five mortality was 38 percent higher than the WFS estimate, suggesting underreporting in the latter. For four surveys (Colombia, Indonesia, Sri Lanka, and Thailand), the DHS estimates were lower than the WFS estimates by 10 percent or more. A more detailed analysis indicated that, in each of these surveys, the percent shortfall in the DHS estimate for child mortality was greater than for infant mortality. This is unusual, since underreporting is more commonly associated with younger age at death.

6.2 RECOMMENDATIONS

The collection of child survival data by means of retrospective surveys will always be subject to reporting error. However, it is possible to develop survey procedures which minimize such errors. Changes in some of the DHS survey procedures would help to realize this goal. These include modifications in the DHS questionnaires, improved training for interviewers, improved instruction manuals, and mandatory follow-up visits to selected respondents by field editors.

Failure to record year of birth information for dead children is a serious lapse in DHS procedures by interviewers and field editors. The necessity of recording these data must be emphasized during interviewer training and should be rigorously checked during field editing. A mechanism to facilitate this process is the inclusion of a box at the end of the birth history section of the questionnaire where the interviewer would be required to check that the year of birth was recorded for dead children. The same procedure would also be helpful as an aid in interviewer training.

When considering the problem of birth displacement, two factors must be taken into account. First, the solution must be sought within the framework of all births in the sixth calendar year before the survey. Given the size of DHS surveys, the number of such births is typically between 500 and 1000. Second, regarding births of children who died, the transferred cases are typically less than 15 percent of the 50 to 100 nonsurviving births reported in the sixth calendar year preceding a survey. Thus, the number of problem cases is relatively small. In a discussion of birth displacement which appears elsewhere in this volume (Arnold, 1990), it is stressed that field editors should spot-check questionnaires with births occurring in the two years prior to the cut-off date for the collection of health data. Also, as soon as the first 1000 or so questionnaires from a survey are received at the data processing office, it would be worthwhile to tabulate the births by calendar year of birth. If the results indicate that there is a problem of birth displacement, the matter can be taken back to the interviewers immediately, to determine the source of the problem and, if possible, the solution to the problem.

Heaping at 12 months for age at death is an ongoing problem in DHS surveys. Typically, it involves 200-300 cases in African surveys and about half as many in surveys in regions where mortality levels are lower. The first step toward minimizing this problem is to ensure that the phenomenon of heaping is fully explained during interviewer training and that it is covered in the interviewer's manual. While this issue was discussed during interviewer training in most DHS surveys, the fact that it was not described in the interviewer's manual tended to de-emphasize its importance. Second, there should be a box at the end of the birth history section of the questionnaire where the interviewer must make a check to indicate that she probed in the case of deaths reported at 12 months of age. Third, field editors should be required to follow up these deaths. Finally, in DHS survey reports consideration should be given to presenting mortality estimates adjusted for heaping of deaths at 12 months of age. Separate mortality estimates might be presented based on (1) the reported data and (2) the data adjusted for heaping of deaths when such an adjustment would result in an increase in the infant mortality rate greater than a specified amount (e.g., 5 percent). It is clear, however, that estimates of mortality from birth to age two or age five are unaffected by heaping of deaths at 12 months of age.

References

Arnold, Fred. 1990. Assessment of the Quality of Birth History Data in the Demographic and Health Surveys. In *An Assessment of DHS-I Data Quality*, Institute for Resource Development. DHS Methodological Reports, No. 1. Columbia, Maryland.

Coale, Ansley J. and Paul Demeny. 1966. *Regional Model Life Tables and Stable Populations*. Princeton, New Jersey: Princeton University Press.

Boerma, Ties. 1988. Monitoring and Evaluation of Health Interventions: Age- and Cause-specific Mortality and Morbidity in Childhood. In *Research and Intervention Issues Concerning Infant and Child Mortality and Health*, 195-218. Proceedings of the East Africa Workshop, International Development Research Centre, Manuscript Report 200e. Ottawa, Canada.

Goldman, Noreen, Ansley J. Coale, and Maxine Weinstein. 1979. *The Quality of Data in the Nepal Fertility Survey.* WFS Scientific Reports, No. 6. Voorburg, Netherlands: International Statistical Institute.

Institute for Resource Development. 1989. *Demographic and Health Surveys Data Processing Manual.* Columbia, Maryland.

Rutstein, Shea Oscar. 1985. Assessment of the Quality of WFS Data for Direct Estimation of Childhood Mortality. In *Assessment of the Quality of Data in 41 WFS Surveys: A Comparative Approach*, by Noreen Goldman, Shea Oscar Rutstein, and Susheela Singh, 63-79. WFS Comparative Studies, No. 44. Voorburg, Netherlands: International Statistical Institute.

Thapa, Shyam and Robert D. Retherford. 1982. Infant Mortality Estimates Based on the 1976 Nepal Fertility Survey. *Population Studies* 36(3):61-80.

Appendix: Birth History Section of the DHS-I Questionnaire

SECTION 2. REPRODUCTION

NO.	QUESTIONS AND FILTERS	CODING CATEGORIES	SKIP TO
201	Now I would like to ask about all the births you have had during your life. Have you ever given birth?	YES.....1 NO.....2	206
202	Do you have any sons or daughters you have given birth to who are now living with you?	YES.....1 NO.....2	204
203	How many sons live with you? And how many daughters live with you? IF NONE ENTER '00'.	SONS AT HOME..... [ ][ ] DAUGHTERS AT HOME..... [ ][ ]	
204	Do you have any sons or daughters you have given birth to who are alive but do not live with you?	YES.....1 NO.....2	206
205	How many sons are alive but do not live with you? And how many daughters are alive but do not live with you? IF NONE ENTER '00'.	SONS ELSEWHERE..... [ ][ ] DAUGHTERS ELSEWHERE..... [ ][ ]	
206	Have you ever given birth to a boy or a girl who was born alive but later died? IF NO, PROBE: Any (other) boy or girl who cried or showed any sign of life but only survived a few hours or days?	YES.....1 NO.....2	208
207	How many boys have died? And how many girls have died? IF NONE ENTER '00'.	BOYS DEAD..... [ ][ ] GIRLS DEAD..... [ ][ ]	
208	SUM ANSWERS TO 203, 205, AND 207, AND ENTER TOTAL. IF NONE ENTER '00'.	TOTAL..... [ ][ ]	
209	CHECK 208: Just to make sure that I have this right: you have had in TOTAL ____ live births during your life. Is that correct? YES [ ] ↓ NO [ ] → PROBE AND CORRECT 201-209 AS NECESSARY		
210	CHECK 208: ONE OR MORE BIRTHS [ ] ↓ NO BIRTHS [ ]		221

211 Now I would like to talk to you about all of your births, whether still alive or not, starting with the first one you had. (RECORD NAMES OF ALL THE BIRTHS IN 212. RECORD TWINS ON SEPARATE LINES AND MARK WITH A BRACKET.)

212 What name was given to your (first, next) baby?	213 Is (NAME) a boy or a girl?	214 In what month and year was (NAME) born? PROBE: What is his/her birthday? OR: In what season?	215 Is (NAME) still alive?	216 IF DEAD: How old was (NAME) when he/she died? RECORD DAYS IF LESS THAN ONE MONTH, MONTHS IF LESS THAN TWO YEARS, OR YEARS.	217 IF ALIVE: How old was (NAME) at his/her last birthday? RECORD AGE IN COMPLETED YEARS.	218 IF ALIVE: Is he/she living with you?
01 (NAME)	BOY 1 GIRL 2	MONTH... YEAR....	YES 1 NO 2 → (GO TO 217)	DAYS.....1 MONTHS...2 YEARS....3 (GO TO NEXT BIRTH)	AGE IN YEARS..	YES 1 NO 2
02 (NAME)	BOY 1 GIRL 2	MONTH... YEAR....	YES 1 NO 2 → (GO TO 217)	DAYS.....1 MONTHS...2 YEARS....3 (GO TO NEXT BIRTH)	AGE IN YEARS..	YES 1 NO 2
03 (NAME)	BOY 1 GIRL 2	MONTH... YEAR....	YES 1 NO 2 → (GO TO 217)	DAYS.....1 MONTHS...2 YEARS....3 (GO TO NEXT BIRTH)	AGE IN YEARS..	YES 1 NO 2
04 (NAME)	BOY 1 GIRL 2	MONTH... YEAR....	YES 1 NO 2 → (GO TO 217)	DAYS.....1 MONTHS...2 YEARS....3 (GO TO NEXT BIRTH)	AGE IN YEARS..	YES 1 NO 2
05 (NAME)	BOY 1 GIRL 2	MONTH... YEAR....	YES 1 NO 2 → (GO TO 217)	DAYS.....1 MONTHS...2 YEARS....3 (GO TO NEXT BIRTH)	AGE IN YEARS..	YES 1 NO 2
06 (NAME)	BOY 1 GIRL 2	MONTH... YEAR....	YES 1 NO 2 → (GO TO 217)	DAYS.....1 MONTHS...2 YEARS....3 (GO TO NEXT BIRTH)	AGE IN YEARS..	YES 1 NO 2

Summary of DHS-I Surveys 1985-1990

Summary of DHS-I surveys, 1985-1990

Region and Country	Date of Fieldwork	Implementing Organization	Respondents	Sample Size	Supplemental Studies, Modules, and Additional Questions
AFRICA					
Botswana	Aug-Dec 1988	Central Statistics Office	All women 15-49	4,368	AIDS, PC, adolescent fertility
Burundi	Apr-Jul 1987	Département de la Population, Ministère de l'Intérieur	All women 15-49	3,970	AM, SAI, adult mortality
Burundi (Husband Survey)	Apr-Jul 1987	Département de la Population, Ministère de l'Intérieur	Husbands	542	KAP study
Ghana*	Feb-May 1988	Ghana Statistical Service	All women 15-49	4,488	AM, SM, WE
Kenya**	Dec-May 1988/89	National Council for Population and Development	All women 15-49	7,150	H
Liberia	Feb-Jul 1986	Bureau of Statistics, Ministry of Planning and Economic Affairs	All women 15-49	5,239	H, TBH, employment status
Mali	Mar-Aug 1987	Institut du Sahel, USED/CERPOD	All women 15-49	3,200	AM, VC, childhood physical handicaps
Mali (Male Survey)	Mar-Aug 1987	Institut du Sahel, USED/CERPOD	Men 20-55	970	KAP study
Ondo State, Nigeria	Sep-Jan 1986/87	Ministry of Health, Ondo State	All women 15-49	4,213	AM, H, TBH
Senegal	Apr-Jul 1986	Direction de la Statistique Ministère de l'Economie et des Finances	All women 15-49	4,415	AM, CD
Sudan	Nov-May 1989/90	Department of Statistics Ministry of Economic and National Planning	EMW 15-49	5,860	H, M, MM, female circumcision, family planning services
Togo	Jun-Nov 1988	Unité de Recherche Démographique Université du Benin	All women 15-49	3,360	AM, H, SAI, marriage history
Uganda	Sep-Feb 1988/89	Ministry of Health	All women 15-49	4,730	AM, H, SAI
Zimbabwe	Sep-Jan 1988/89	Central Statistical Office	All women 15-49	4,201	AIDS, AM, H, PC, SAI, WE
ASIA/NEAR EAST/NORTH AFRICA					
Egypt	Oct-Jan 1988/89	National Population Council	EMW 15-49	8,911	AM, CD, H, MM, PC, SAI, WE, women's status
Indonesia	Sep-Dec 1987	Central Bureau of Statistics National Family Planning Coordinating Board	EMW 15-49	11,844	PC, SM
Morocco	May-Jul 1987	Ministère de la Santé Publique	EMW 15-49	5,982	AM, CD, H, S
Nepal (In-depth)	Feb-Apr 1987	New Era	CMW 15-49	1,623	KAP–gap survey
Sri Lanka	Jan-Mar 1987	Dept. of Census and Statistics Ministry of Plan Implementation	EMW 15-49	5,865	AM, H, NFP
Thailand	Mar-Jun 1987	Institute of Population Studies Chulalongkorn University	EMW 15-49	6,775	AM, S, SAI
Tunisia	Jun-Oct 1988	Office National de la Famille et de la Population	EMW 15-49	4,184	AM, CD, H, S, SAI

Region and Country	Date of Fieldwork	Implementing Organization	Respondents	Sample Size	Supplemental Studies, Modules, and Additional Questions
LATIN AMERICA & CARIBBEAN					
Bolivia	Mar-Jun 1989	Instituto Nacional de Estadística	All women 15-49	7,923	AM, CD, H, MM, PC, S, WE
Bolivia (In-depth)	Mar-Jun 1989	Instituto Nacional de Estadística	All women 15-49	7,923	Health
Brazil	May-Aug 1986	Sociedade Civil Bem-Estar Familiar no Brasil	All women 15-44	5,892	AM, H, PC, SM, abortion, young adult use of contraception
Colombia	Oct-Dec 1986	Corporación Centro Regional de Población Ministerio de Salud	All women 15-49	5,329	AM, PC, SAI, SM
Dominican Republic	Sep-Dec 1986	Consejo Nacional de Población y Familia	All women 15-49	7,649	NFP, S, SAI, SM family planning communication
Dominican Rep. (Experimental)	Sep-Dec 1986	Consejo Nacional de Población y Familia	All women 15-49	3,885	
Ecuador	Jan-Mar 1987	Centro de Estudios de Población y Paternidad Responsable	All women 15-49	4,713	SAI, CD, H, employment
El Salvador	May-Jun 1985	Asociación Demográfica Salvadoreña	All women 15-49	5,207	S, TBH
Guatemala	Oct-Dec 1987	Instituto de Nutrición de Centro América y Panamá	All women 15-44	5,160	H, S, SAI
Mexico	Feb-May 1987	Dirección General de Planificación Familiar Secretaría de Salud	All women 15-49	9,310	H, NFP, S, employment
Peru	Sep-Dec 1986	Instituto Nacional de Estadística	All women 15-49	4,999	H, NFP, employment, cost of family planning
Peru (Experimental)	Sep-Dec 1986	Instituto Nacional de Estadística	All women 15-49	2,534	
Trinidad and Tobago	May-Aug 1987	Family Planning Association of Trinidad and Tobago	All women 15-49	3,806	AM, NFP, breastfeeding

CMW = currently married women
EMW = ever-married women

AIDS = acquired immune deficiency syndrome
AM = anthropometric measurements
CD = causes of death (verbal reports of symptoms)
H = additional health questions
M = migration
MM = maternal mortality
NFP = natural family planning
PC = pill compliance
S = sterilization
SAI = service availability information
SM = social marketing
TBH = truncated birth history
VC = value of children
WE = women's employment

* Data available for 943 husbands interviewed with a husband's questionnaire
** Data available for 1,133 husbands interviewed with a husband's questionnaire