USAID
FROM THE AMERICAN PEOPLE

# EVIDENCE OF OMISSION AND DISPLACEMENT IN DHS BIRTH HISTORIES

## DHS METHODOLOGICAL REPORTS 11

**SEPTEMBER 2014**

This publication was produced for review by the United States Agency for International Development. It was prepared by Thomas W. Pullum of ICF International and Stan Becker of Johns Hopkins University.

DHS Methodological Reports No. 11

# Evidence of Omission and Displacement in DHS Birth Histories

Thomas W. Pullum[1]
Stan Becker[2]

ICF International
Rockville, Maryland, USA

September 2014

[1] The DHS Program
[2] Bloomberg School of Public Health, Johns Hopkins University; sbecker@jhsph.edu

*Corresponding author:* Thomas W. Pullum, International Health and Development, ICF International, 530 Gaither Road, Suite 500, Rockville, MD 20850, USA; phone: 301–572–0950; fax: 301–572–0999; email: tom.pullum@icfi.com

**Acknowledgments:** The authors are grateful for helpful comments, at various stages of preparation, from Trevor Croft and Jeremiah Sullivan, and for valuable reviews by Stephane Helleringer and Shea Rutstein.

Editor: Bryant Robey
Document Production: Yuan Cheng

This study was carried out with support provided by the United States Agency for International Development (USAID) through the MEASURE DHS project (#GPO-C-00-08-000080-00). The views expressed are those of the authors and do not necessarily reflect the views of USAID or the United States Government.

The DHS Program assists countries worldwide in the collection and use of data to monitor and evaluate population, health, and nutrition programs. For additional information about The DHS Program contact: DHS Program, ICF International, 530 Gaither Road, Suite 500, Rockville, MD 20850, USA; phone: 301-407-6500, fax: 301-407-6501, email: reports@dhsprogram.com, Internet: www.dhsprogram.com.

Recommended citation:

Pullum, Thomas W., and Stan Becker. 2014. Evidence of Omission and Displacement in DHS Birth Histories. DHS Methodological Reports No. 11. Rockville, Maryland, USA: ICF International

# Contents

# Tables

# Figures

# Preface

The Demographic and Health Surveys (DHS) Program is one of the principal sources of international data on fertility, family planning, maternal and child health, nutrition, mortality, environmental health, HIV/AIDS, malaria, and provision of health services.

One of the objectives of The DHS Program is to continually assess and improve the methodology and procedures used to carry out national surveys as well as to offer additional tools for analysis. Improvements in survey methods will enhance the accuracy and depth of information relied on by policymakers and program managers in low- and middle-income countries.

Although data quality is a focus of the DHS Methodological Reports series, the reports also examine issues of sampling, questionnaire comparability, survey procedures, and methodological approaches. The topics explored in this series are selected by The DHS Program in consultation with the US Agency for International Development.

It is hoped that the DHS Methodological Reports will be useful to researchers, policymakers, and survey specialists, particularly those engaged in work in low- and middle-income countries, and will be used to enhance the quality and analysis of survey data.

Sunita Kishor

Director, The DHS Program

# Abstract

This report assesses the quality of the birth history data in 192 DHS surveys conducted since 1990. The birth histories are the source of the age-specific fertility rates, the total fertility rate (TFR), and the under-5 mortality rates that are widely used to describe levels, differentials, and trends in fertility and child mortality in developing countries. The assessment is based on three criteria to identify potential omission of births and/or deaths and three criteria to identify potential displacement, or misreporting, of date of birth, age at death, or both. Extreme values should be viewed as symptoms rather than a conclusive diagnosis of poor data quality. The birth histories in most DHS surveys appear to be of excellent quality, although several surveys do show symptoms of omission, displacement, or both. With the selected indicators and criteria, omission of births is usually at the level of 2 percent or less, and only rarely exceeds 5 percent. Displacement of births is also usually 2 percent or less, and only rarely exceeds 3 percent of births in the past 10 years. The omission of deaths is a higher percentage; but usually less than 5 percent, although it may exceed 10 percent in some surveys. Displacement of deaths is usually less than 5 percent, and almost never more than 10 percent. With some unevenness across regions, DHS-6 appears to be the phase of DHS that had the lowest overall levels of incomplete birthdates, flagged dates of death, and omission and displacement of births and deaths.

**KEY WORDS:** data quality, birth history, fertility rates, under-5 mortality rates

# Executive Summary

The Demographic and Health Surveys Program strives to achieve and maintain the highest possible quality of basic data. As part of this effort, The DHS Program performs detailed field checks and supervision during the data collection process, carefully edits the data for internal consistency during data processing, includes indicators of completeness and consistency within every data set, provides several tables on data quality within every main report, and periodically undertakes comprehensive reviews of large numbers of surveys. This report reviews the birth histories in 192 DHS surveys conducted between 1990 and 2013, and births during the 10 years prior to the interviews in each survey. Its specific objective is to assess the data in the birth histories for evidence of omission or displacement of births and deaths.

The birth histories are lists of all the live births ever born to the main survey respondents, women age 15-49 at the date of interview. The lists include the month and year of each birth, the sex of the child, and whether it was part of a multiple birth. If the child is no longer alive at the date of interview, the age at death is provided. This basic information, when combined with the woman's month and year of birth, and the month and year of the interview, is all that is required to estimate age-specific fertility rates, the total fertility rate (TFR), and the under-5 mortality rates, which are among the most important results of DHS surveys.

The data in the birth histories are subject to reporting errors that have potential impact on the fertility and mortality estimates. Sometimes the birth *dates are incomplete or are clearly incompatible with one another. Sometimes the age at death is inconsistent with other age- or date-related information. During data processing, each birth is given a code for completeness of the date reporting and each death is given a code or flag for inconsistency related to the age at death. In addition to those two indicators, this report examines six other potential indicators of data quality—three to identify omission and three to identify displacement, a term that refers to incorrect reporting of date of birth or age at death.

Systematic omission of births will cause the fertility rates to be artificially low. If the omissions are primarily of children who died, then the under-5 mortality rates will also be artificially low. Such downward biases in the rates will have serious implications. If the probabilities of omission are greater in some subpopulations than others, then analyses of differentials may be misleading; if the probabilities are greater in one survey than in another in the same country, then analyses of trends may be misleading. Omission is thus very important, arguably more important than displacement, but it is much harder to identify.

In this report, omission is indicated by a deviation, in some characteristic of the survey's birth history data, from an expected value based on international evidence. The first indicator of omission used here is the deviation of the sex ratio at birth, or SRB, from an expected value. The second indicator is the deviation of the sex ratio of neonatal deaths (deaths in the first month) from an expected value. The third indicator is the deviation of the proportion of infant deaths that are neonatal from an expected value that is based on the ratio of infant deaths to births.

We also include three indicators of displacement, based on expected regularities within the data rather than reference to international evidence. The first indicator is heaping on reported age at death at 12 months. The second indicator is heaping on reported age at death at 7 days. The third indicator of displacement measures potential transfers of birth dates, from the first calendar year inside the interval for the child health questions, backward into the last calendar year before the interval.

The six indicators are augmented with measures of how many births or deaths would have to be added or shifted to bring a survey into exact compliance with all the criteria. This is achieved with an iterative reweighting procedure that effectively adds or shifts the births and dates by adjusting the sampling weights.

For each survey, the results are presented in terms of the values of the indicators and the implied impact on the numbers of births and deaths. Surveys at the extremes are listed, and averages for each region and phase of DHS surveys are also provided.

The results are sensitive to the selection of criteria and the reference values for the omission criteria. With alternative criteria and reference values, the results would be different, especially in terms of the identification of specific surveys with relatively high levels of omission. High values of the indicators should be viewed as symptoms rather than a conclusive diagnosis of poor data quality. More in-depth analysis of the birth histories is recommended for surveys with high values.

The birth histories in most DHS surveys appear to be of excellent quality, although some surveys do show symptoms of omission, displacement, or both. With the selected indicators and criteria, omission of births is usually at the level of 2 percent or less. Displacement of births is also usually 2 percent or less. Omission and displacement of births are less likely in DHS-6 than previously.

The omission of deaths, i.e., of births that resulted in deaths, is higher and may exceed 10 percent in some surveys. Displacement of deaths may exceed 10 percent in some surveys, but like displacement of births, it appears to be less of an issue in recent surveys.

This report on data quality is one of a group of three being produced at the same time. Methodological Report #12, by Bruno Schoumaker, examines the fertility rates themselves and Methodological Report #13, by Saifuddin Ahmed et al., examines the quality of the sibling histories and the estimates of adult and maternal mortality.

# 1. Introduction and Overview

DHS surveys are the main source of estimates of levels, differentials, and trends in fertility and under-5 mortality in developing countries. Births recorded in the retrospective birth histories permit direct estimation of age-specific fertility rates and the total fertility rate (TFR). Information about the survivorship of children and when they died (if they died) also permit direct estimation of neonatal, postneonatal, infant, child, and under-5 mortality rates. When combined with relatively recent statistical methods and software, the analysis of such data can include confidence intervals, hypothesis tests, covariates, multi-level modeling, projections, and other sophisticated approaches to the analysis of fertility and mortality. The value of these measures and analyses depends on the underlying accuracy of the events and the dates of those events in the birth histories.

This report is part of an ongoing effort to assess, monitor, and improve the quality of DHS data. Its specific objective is to assess the data in the birth histories for evidence of omission or displacement of births or deaths. This objective could be approached from at least three different perspectives. One possibility would be to examine the process of collecting the data. For example, the responses could be linked to data that have been collected as part of a Health and Demographic Surveillance System (HDSS), as Helleringer et al. (2014) describe for similar data on sibling survival. Another field-based approach to data quality would be to re-interview the same respondents, checking for consistency of responses, and reconciling discrepancies. A second perspective, which is applied in a companion report by Schoumaker (2014), is based on the calculated rates, examining the consistency between successive surveys or the consistency with rates calculated from other sources. A third perspective, intermediate to the other two—and the one used here—is based on the internal consistency of the recoded data files of births, after the data have been collected and prior to the construction of rates.

These perspectives on the quality of the data on births and deaths are complementary and inter-related. Some errors that occur during fieldwork may be conspicuous but actually have virtually no impact on the calculated rates. Such errors should be minimized, if their causes can be identified, but their impact may be small. For example, rounding or heaping of ages or dates during data collection may be a symptom of other problems, but a moderate level of unbiased rounding will tend to cancel out and have virtually no impact on most rates. However, if such rounding and heaping occurs at boundaries for inclusion or exclusion, it can produce bias. Thus, child deaths rounded at 12 completed months of age, just outside the range for infant deaths, will bias the infant mortality rate downward, but rounding at 6 months or 18 months will have no effect on the standard rates. For another example, if children who died early—in the first day, week, month, or year—tend to be omitted, the ramifications for the estimates of child mortality can be dramatic, but the fertility estimates will suffer only slightly, because such deaths are only a small fraction of all births.

As mentioned, this report will focus on the birth histories in single surveys, a micro-level perspective that is intermediate to the data collection and the calculation of summary rates. There will be some references to fieldwork but we will not, for example, include data from field checks. There will be references to rates, but only approximations to rates for which the usual calculations are rather complex.

As a matter of general policy, DHS does not adjust data to compensate for evidence of deficiencies such as omission and displacement. Some minor exceptions to this policy can be identified—for example, imputation of some dates that are incomplete or inconsistent, and consolidation of highly unlikely measurements of height and weight under "implausible" codes. But for the most part, driven by the general desire to maximize the analytical possibilities for researchers outside of DHS, the data files are distributed without adjustments. When DHS produces indicators for the main reports on surveys or even for comparative reports, etc., there is again, generally, no adjustment for data-quality issues. If DHS data files or reports did include adjustment, then questions would inevitably arise about the impact of the adjustments, how they were made, whether they could have been made differently, and how they can be undone. The general policy is thus to hand over to users the option of whether and how to make adjustments. In that spirit, this report will not explicitly provide adjusted estimates of fertility and mortality rates.

Births are subject to two principal types of errors: omission and displacement. The omission of a birth is the more serious type of error, and also is much harder to detect than displacement. This report will consider three possible types of omission: a tendency for births of one sex to be systematically under-reported; a tendency for neonatal deaths of one sex to be systematically under-reported; and a tendency for children who died within the first month—neonatal deaths—to be systematically under-reported, relative to deaths later in the first year. The levels of these three types of omission are estimated by comparing the observed sex ratio of births, the sex ratio of neonatal deaths, and the proportion of infant deaths that are neonatal with expectations from international data known to be of high quality.

The most serious type of omission may well be omission of children who died before the survey with a pattern that depends little, if at all, on the age when the child died. We will discuss this type of omission, but are unable to provide methods to identify it with a single survey. We will also not attempt to identify omission that results from possible omission or nonresponse of women from the survey. This type of omission is believed to be rare, but if it occurs it is probably not random with respect to the woman's fertility and the survival of her children.

Displacement in the birth history refers to an error in the stated date of an event. Three types of displacement will be considered. Two of these could be described as heaping, a tendency for events to be shifted toward a central date of birth or death or age at death. Heaping of age at death at exactly 7 days or exactly 12 months will be singled out because it affects the boundaries of the age intervals for early neonatal and infant deaths. The third type of displacement is potential transfers of birth dates across the boundary for the questions on child health. Many surveys show a dip in the numbers of births reported to be just inside the window or interval for the child health questions, and a corresponding bump in the numbers just preceding the window. The likely explanation is that some interviewers tend to reduce their workload by shifting some births back in time. This type of displacement is important because it can distort evidence of recent trends in fertility. The levels of the three types of displacement are estimated by comparing the observed distributions of ages at death and dates of births with a hypothetical uniform or smooth distribution in a range of ages at death and dates of birth.

Possible omission and displacement are not limited to the types listed previously. In many surveys it is obvious, for example, that birth years have been estimated in two steps: by first estimating the current age of a surviving child in completed years, and then subtracting that age from the year of the interview. Estimates of a child's age may be heaped at even numbers, and as a result the birth histories may show a regular sequence of dips and bumps in alternating years. In some surveys the birth year has been based on identification cards, on which the birth year had been calculated from the estimated age of the child when the card was issued, with heaping on certain ages. In some countries there may a tendency to bias a child's age upward (or perhaps downward), leading to spurious evidence of recent fertility declines (or increases). For example, Pullum and Stokes (1997) developed an approach to the well-documented tendency for children in Pakistan to be reported in surveys and censuses as being older than they actually are. There is a tendency in some surveys for some years of birth, such as the year 2000, to be over-reported. The strategy used in this report could be extended to include checks for such additional types of displacement, as well as additional types of omission.

Indices are produced for each survey to describe the level of each type of omission or displacement. The birth histories are then adjusted in such a way that after adjustment all six indices have the value zero. That is, an adjustment procedure is used that forces the data to satisfy all of the criteria. The net change after these adjustments will serve as an overall indicator of data quality: the lower the magnitude of change or adjustment, the better the survey. In this report, we do not recalculate the rates after adjustment and compare with the rates before adjustment, partly because of the general policy against adjustment, as mentioned above, and partly because we do not want to give the mistaken impression that these three types of omission and three types of displacement comprise a full inventory of possible data problems.

The procedure to correct[1] the birth histories for possible omission and displacement does not involve adding births or adding deaths or altering dates. That strategy might initially seem attractive, but it would be virtually impossible to implement without introducing even more serious inconsistencies, such as births that are too closely spaced. Instead, the sampling weights will be adjusted. To compensate for omission of births or deaths, we increase the weights[2] of births or deaths[3] that are believed to be under-represented. To compensate for displacement, we increase the weights for the under-represented category and decrease the weights for the over-represented category, in such a way that the sum of the weights for the two categories is maintained. The adjustments can be made for one type of omission or displacement at a time, or for all in combination, by looping through the six types of omission or displacement sequentially and repeatedly until convergence is reached.

Reweighting is believed to be new in this context, but it has a long history in statistics, for example in the form of post-stratification reweighting of census data (Little 1993) and the iterative adjustment of tables to match specified marginal criteria (Deming 1943, chapter vii). Iteratively reweighted least squares (IRLS) is a commonly used computing technique for maximum-likelihood estimation or other kinds of optimization (Basu 2005).

Analyses of data quality often do not use weights at all (see, for example, Pullum 2006). In this report some of the results will be unweighted, but most will be weighted (or reweighted) as just described. It is appropriate to use weights to assess the impact on estimates of fertility and child mortality rates, because DHS always uses weights when calculating those rates.

This report does not attempt to explain variations in omission and displacement, but the adjustment procedure would allow a statistical analysis of such variation. For every birth in the data file, the combined effect of the adjustments can be expressed in terms of the net change in the weight. Then, for example, the absolute value of the log of the ratio of the two weights could provide a metric for the magnitude of change—a number that is 0 if there is no change and positive if there is any change, either an increase or a decrease in the weight. Variation in this metric across characteristics of the respondent, interviewer ID code, successive surveys, different countries and regions, etc., could be assessed, even though such analysis is beyond the scope of this report.

---

[1] Although the word "correct" will be used occasionally in this report, as a verb or an adjective, it must be understood that modifications to the original birth histories are only intended to improve the data, and only at an aggregated level. It is virtually impossible to identify an error in an individual birth history, let alone to remove that error. Some other terms, such as "error," should also be interpreted loosely.

[2] The weight variable in DHS data files is indeed a sampling weight. When these weights are adjusted upward or downward they could be described as "analytical weights," for example, to distinguish them from the original sampling weights.

[3] The sampling weights apply to specific cases in the data, such as households, women, or children. Strictly speaking, they do not apply to an outcome, such as a birth or a death. This distinction will be clarified below.

The point will be made repeatedly, but it must be emphasized that a deviation from an expected value is not equivalent to an error. The criterion or reference value may simply not apply, particularly for evidence of omission. For example, the sex ratio of births in the population may genuinely differ from a criterion value that is derived from some external source. Moreover, deviations are always expected just because of sampling variability. Measurement error, the component that we wish to assess, is inevitably confounded with incorrect specification of the criterion and with sampling error.

This report is one of a series on data quality issues. It is most closely related to early reports on the birth histories by Arnold (1990) and on child survival by Sullivan, Bisego, and Rutstein (1990), both on DHS-1; a report on DHS data quality by Marckwardt and Rutstein (1996), comparing DHS-1 and DHS-2; a report on the quality of age and date reporting for surveys conducted from 1985 to 2003, i.e., DHS-1 through DHS-4 (Pullum 2006), and the report by Schoumaker (2014), appearing at the same time as this one.

# 2. Data

The analysis includes 192 surveys conducted by DHS from 1990 through 2013, virtually all of the standard surveys ever conducted, except for surveys in the first phase of DHS (1985-89), which were examined by Arnold (1990), or the most recent surveys in the sixth phase of DHS that had not yet been released when the final selection had to be made. The phases correspond with calendar years as follows: phase 2: 1990-1993; phase 3: 1994-1998; phase 4: 1999-2004; phase 5: 2005-2009; phase 6: 2010-2013. Malaria Indicator Surveys (MIS), AIDS Indicator Surveys (AIS), and interim, experimental, or restricted surveys are not included. The Peru Continuous Survey (PCS), which began at the end of 2003, is included with files for 2004-2006, 2007-2008, and separate single-year files for 2010 and 2011.[4] Countries will be classified by region, using the same categorization as STAT Compiler (www.STATcompiler.com). Because the greatest share of surveys has been in sub-Saharan Africa, that region is partitioned into sub-regions. Table 2.1 shows the number of surveys in each region during the successive phases of DHS. A complete list of the surveys, with numbers of births, is provided in the Appendix. In all surveys, births more than 10 years before the month of interview are ignored.[5]

**Table 2.1. The number of surveys in this report in each combination of geographic region and phase of DHS**

| Region | 2 | 3 | 4 | 5 | 6 | Total |
|---|---|---|---|---|---|---|
| Western Africa | 4 | 9 | 9 | 8 | 7 | 37 |
| Middle Africa | 1 | 3 | 3 | 3 | 3 | 13 |
| Eastern Africa | 5 | 9 | 12 | 7 | 8 | 41 |
| Southern Africa | 1 | 1 | 2 | 2 | 1 | 7 |
| N Africa / W Asia / Europe | 4 | 3 | 6 | 8 | 2 | 23 |
| Central Asia | 0 | 3 | 1 | 0 | 1 | 5 |
| South and SE Asia | 3 | 8 | 8 | 8 | 5 | 32 |
| Latin America and Carib. | 5 | 9 | 7 | 8 | 5 | 34 |
| Total | 23 | 45 | 48 | 44 | 32 | 192 |

The largest surveys, by far, are the three surveys of India carried out in 1992-93, 1998-99, and 2005-06. The weighted numbers of births and deaths in the past 10 years in those surveys are as follows:

- India 1992-93 DHS: 128,372 births and 12,764 deaths
- India 1998-99 DHS: 120,082 births and 11,150 deaths
- India 2005-06 DHS: 109,384 births and 8,553 deaths

[4] The archived DHS files for the years of the PCS are occasionally regrouped. On the date when the files for this report were downloaded, the 2009 data were not in the archive, an omission that was only recognized after the analysis had been completed. It is not clear how the PCS should contribute to the “number of surveys.” We count it as four surveys simply because we are using four files, for seven of the eight years 2004 to 2011. A more detailed assessment of the PCS would include separate files for each year.

[5] This analysis includes births in the month of interview, as well as months 1-120 before the interview. The standard DHS rates do not include births and deaths in the month of interview.

Apart from those three surveys, the weighted number of births ranged from 1,260 to 55,594, and the number of deaths ranged from 47 to 8,087. The mean and median numbers of births were 15,112 and 13,539, respectively. The mean and median numbers of deaths were 1,417 and 1,244, respectively. We retained the surveys with relatively few births or deaths in the past 10 years, but sampling error is clearly a greater issue for proportions, ratios, and rates calculated from those surveys, compared with larger surveys.

Birth histories are obtained during the interviews of eligible women respondents, generally women age 15-49 who are de facto residents of the sampled household at the date of interview. Some surveys are limited to ever-married women.

During the interview, the information about births is obtained in a series of steps. First, prior to the detailed birth history, the woman is asked a series of questions about whether she has had any children, how many, how many boys or girls, and separately for the boys and girls, how many are alive and living with the mother, how many are alive but not living with the mother, and how many have died. These numbers of boys and girls, alive and dead, are revised during the interview if they do not match with the birth history after it has been collected, but there is not a record of whether any such reconciliation actually happened in the field.

If the woman reports at least one live birth, then the birth history is recorded, with one row for each birth, beginning with the earliest birth.[6] The questions are listed below, using generic question numbers.

> Q211. "Now I would like to record the names of all your births, whether still alive or not, starting with the first one you had."
>
> Q212. "What name was given to your [first/next] baby?"
>
> Q213. "Is [NAME] a boy or a girl?"
>
> Q214. "Were any of these births twins?" (Note: each child in a multiple birth is listed separately)
>
> Q215. "In what month and year was [NAME] born?"
>
> Q216. "Is [NAME] still alive?"
>
> Q217. "If alive: How old was [NAME] at his/her last birthday?"
>
> Q218. "If alive: Is [NAME] living with you?"
>
> Q219. If alive: interviewer records the household line number of the child
>
> Q220. If dead: "How old was [NAME] when he/she died" (Note: see detail below)
>
> Q221. "Were there any other live births between [NAME OF PREVIOUS BIRTH] and [NAME], including any children who died after birth?"

At the beginning of the listing with Q211, and after each child with Q221, the importance of including children who have died is stated explicitly. If, during the collection of the birth history, another child is recalled, then that child is to be inserted in the history, but there is no record of how often this happens.

It is important to distinguish birth histories, as collected by DHS, from pregnancy histories. A pregnancy history would be a list of all pregnancies that the woman recalls, with any possible pregnancy termination, not just a live birth.[7] The alternative terminations would be a spontaneous abortion or miscarriage, an

---

[6] The birth order of the child is coded in the data file as "bord," with "bord=1" for the earliest birth. Another variable, "bidx," numbers children born in the in the child health window in reverse order, beginning with "bidx=1" for the most recent birth.

[7] A handful of DHS surveys collected pregnancy histories, but we only use the births file, as with all other surveys.

induced abortion, or a stillbirth. DHS normally supplements the birth history with the so-called "calendar," a month-by-month listing of pregnancy status and contraceptive use during the five years before the interview. The calendar includes all possible pregnancy terminations, but it is collected mainly for information about contraceptive use dynamics—that is, method failure, discontinuation, or switching. There is some reconciliation of the calendar with the birth history, during fieldwork and during data processing. This report will not refer to any of the calendar data.

Note that Q215 (month and year of birth) and Q217 (current age of the child, in complete years, if still alive) include some redundancy. If we know when the child was born (and the date of the interview, which of course is always recorded), then current age in years could be calculated. There are two reasons for explicitly asking both questions. First, if one of them is missing, then the other can be a substitute; second, if they disagree, then we can infer that there is a data-quality issue. It is likely that there is some supplementation and reconciliation of these two items in the field, but there is no record of that.

In the data processing of this information, the name of the child is of course not retained. The responses are coded in the "b variables," as listed below, and are included on the mother's computer record with the index or subscript "bidx", which is 1 for the most recent birth. The woman's file is referred to as the IR ("individual recode") file. All the live births are also listed in a BR ("births recode") file with one record per birth, indexed by the mother's id code and bidx, as well as the child's chronological birth order (bord), with no subscripts in the variable names.

Almost all of the analysis in this report is limited to the BR or births file,[8] but some reference will also be made to the IR or women's file because the calculation of age-specific fertility rates requires the birthdates of all women, not just the women who had a birth in a specific interval of time. The only variables required in our analysis, apart from the "all women factor," awfactt,[9] and a code to identify the country and survey, are the following:

- v005: sample weight
- v008: month (cmc) of interview
- v011: month (cmc) of woman's birth
- v017: first month (cmc) in the "health window" (a single survey-specific number)
- b2: calendar year of child's birth (could be calculated from b3)
- b3: month (cmc) of child's birth (may have required some imputation)
- b4: sex of child (1=boy, 2=girl)
- b5: survivorship of child to date of mother's interview (0=died, 1=survived; could be calculated from b6 or b7, because if the child survived, then b6 and b7 have "missing" codes)
- b6: child's age at death (days/months/years)
- b7: child's age at death (in months)

[8] Users of DHS data may be more familiar with the KR file ("kids recode") or child file, which is limited to children born in the child health window. The BR file includes the same data as the KR file, plus the b variables for children born before the child health window, for whom the b variables are the only information. Both the BR and KR files include most of the mother's information in the IR file, attached to the child.

[9] Surveys of ever-married women, rather than all women, require a cluster-specific adjustment to the weights for the calculation of fertility rates. Specifically, for such surveys, v005 must be multiplied by awfactt/100. Similar factors specific for region, education, and wealth grouping are also calculated for such surveys.

b10: completeness of information about date of birth (a code other than 1 indicates incomplete information)

b13: flag for age at death (a code other than 0 indicates inconsistent or incomplete information)

The four variables that provide a date (v008, v011, v017, and b3) are coded in century month codes (cmc), which begin with code 1 for January 1900. If M is the ordinal number of the calendar month (1 to 12) and Y is the calendar year, then an event that occurred in year Y and month M will have cmc=M+12*(Y-1900). For example, January 2000 is assigned cmc=1201. Dates for Ethiopia and Nepal are converted to the Gregorian calendar.

The first month (cmc) of the child health window is typically January for the fifth calendar year preceding the first month and year of fieldwork. For example, if data collection began in October 2010, v017 would normally be January of 2005, cmc=1261. The questionnaire would require the collection of the additional information about immunizations, nutrition, recent illnesses and treatment, etc., for all children born since January 2005.

During data processing, the health information is retained only for children born during the 60 months prior to the month of interview. In this example, the first interviews would have been in October 2010, so the health information collected for children born during January-September of 2005 would always be discarded, and more months would be discarded for interviews conducted later than October 2010. This specification of the health window—almost always to begin in January of a calendar year—usually involves discarding some data, but is justified by practical considerations, such as uncertainty about when the fieldwork will begin and the desirability of having a fixed specification for all interviews.

For some surveys the health window has been shorter than five years. Regardless of the length, there is often evidence of some displacement across the boundary for the health window. The evidence appears in the distribution of births by calendar year, provided in all main survey reports in Table 4 of Appendix C. Displacement shows up as an irregularity in this distribution, with a dip in the number of births reported for the earliest year within the health window and a bump in the number reported for the preceding year. In the example above, it could appear that some births that actually occurred in 2005 were reported for 2004. It could even appear that some births occurring in 2006 were reported for 2004 or earlier.

The Senegal 2010-11 DHS survey will be used to illustrate this pattern. The choice of this survey should not be taken to mean that it was of particularly poor quality. The fieldwork extended from October 2010 to May 2011 (only 10 births are recorded for May 2011; the fieldwork essentially ended in April). Table 2.2 shows the unweighted numbers of births reported in this survey for the calendar years 2000-2011. There is a dip in the number of births reported for 2005 and a bump for 2004, suggesting that perhaps 10 to 12 percent of the births that actually occurred in 2005 were misreported for 2004. The ten-year interval for the births begins in 2000 or 2001, depending on the month of the interviews in 2010 and 2011. Although masked by this cutoff, substantially more births were reported in 2000 than in 1999 or 2001, a pattern of heaping on 2000 that is found in many DHS surveys. About a third of the interviews were conducted in 2010, and because of reduced exposure to 2010 we would expect fewer births that year, compared with, say 2009. The fact that more births are reported for 2010 than for any other year also suggests digit preference.

**Table 2.2. Unweighted frequency distribution of b2, calendar year of birth, in the Senegal 2010-11 DHS Survey; fieldwork for this survey was conducted from October 2010 to May 2011, limited to children who were born during the 10 years (120 months) before the mother's interview**

| b2: Year of birth | Births |
|---|---|
| 2000 | 113[a] |
| 2001 | 1,533[a] |
| 2002 | 2,249 |
| 2003 | 2,325 |
| 2004 | 2,595 |
| 2005 | 2,071 |
| 2006 | 2,247 |
| 2007 | 2,441 |
| 2008 | 2,463 |
| 2009 | 2,475 |
| 2010 | 2,624[b] |
| 2011 | 278[b] |
| Total | 23,414 |

[a] The numbers of births in 2000 and 2001 are reduced because of the ten-year cutoff.

[b] The numbers of births in 2010 and 2011 are reduced because of the dates of fieldwork.

Systematic displacement across the boundary for the health questions is due to the interviewer rather than the respondent. At the point in the interview when the birth history is being collected, the respondent does not know that extra questions will be asked about some children; only the interviewer knows that. However, the respondent has a secondary role; displacement probably is more likely if the respondent is unsure about the age of a child or the date of birth. In such cases it is likely that the interviewer tends to resolve the uncertainty in the direction that will reduce her (women are interviewed by women) workload. Thus we expect more displacement in settings with less accurate knowledge of ages and dates, as well as for interviewers with less thorough training and supervision.

Specification of the woman's (mother's) date of birth, and the child's date of birth, down to the level of a month, often involves some imputation or editing. The month may have to be estimated on the basis of only a specified year of age and/or year of birth. The imputation procedure resolves incompleteness by random assignment of a month within a plausible interval, but even if a child's date of birth is provided by the mother as a calendar year and month, DHS subjects the birth histories to some editing and possible shifting in order to avoid impossibly close spacing of births. It is not believed that any biases are introduced during the editing and imputation process.

If the child died, the mother is not asked for the month of death, but rather the age at death: "How old was (NAME) when he/she died?" These additional instructions are given to the interviewer: "Record days if less than one month; months if less than 2 years; or years." The responses are then recorded with a three-digit code, b6, with first digit 1 for days, 2 for months, and 3 for years. To illustrate this rather unusual type of variable, the unweighted distribution of b6 in the Senegal 2010-11 survey is given in Table 2.3. Like Table 2.2, Table 2.3 is limited to births in the 10 years (120 months) before the month of interview; the data files omit births in the month of interview. The variable b6 is missing ("." in Stata) for the children who were still living at the date of interview. Non-missing codes for b6 are given just for the 1,910 children who died. (One child had code 998, and is not included in Table 2.3 but is included in Table 2.4.) Table 2.3

shows that 803 of these children had their age at death coded in days, with prefix 1; 640 have prefix 2 for months, and 467 have prefix 3 for years.

**Table 2.3. Unweighted frequency distribution of b6, age at death, in the Senegal 2010-11 DHS Survey, includes all children who were born during the 10 years (120 months) before the mother's interview and died before the interview**

| | First digit of b6 | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | Total |
| Second and third digits of b6 | Deaths | Deaths | Deaths | Deaths |
| 00 | 212 | 0 | 0 | 212 |
| 01 | 173 | 89 | 66 | 328 |
| 02 | 48 | 71 | 189 | 308 |
| 03 | 77 | 60 | 86 | 223 |
| 04 | 32 | 40 | 41 | 113 |
| 05 | 37 | 40 | 29 | 106 |
| 06 | 30 | 47 | 20 | 97 |
| 07 | 52 | 30 | 17 | 99 |
| 08 | 40 | 30 | 6 | 76 |
| 09 | 14 | 35 | 2 | 51 |
| 10 | 11 | 12 | 1 | 24 |
| 11 | 2 | 24 | 0 | 26 |
| 12 | 6 | 54 | 6 | 66 |
| 13 | 1 | 16 | 0 | 17 |
| 14 | 6 | 7 | 0 | 13 |
| 15 | 28 | 11 | 0 | 39 |
| 16 | 2 | 9 | 0 | 11 |
| 17 | 1 | 3 | 0 | 4 |
| 18 | 2 | 37 | 3 | 42 |
| 19 | 3 | 2 | 0 | 5 |
| 20 | 6 | 2 | 0 | 8 |
| 21 | 7 | 2 | 0 | 9 |
| 22 | 1 | 3 | 0 | 4 |
| 23 | 1 | 9 | 0 | 10 |
| 24 | 1 | 2 | 1 | 4 |
| 25 | 2 | 1 | 0 | 3 |
| 26 | 0 | 3 | 0 | 3 |
| 28 | 1 | 1 | 0 | 2 |
| 30 | 5 | 0 | 0 | 5 |
| 36 | 1 | 0 | 0 | 1 |
| 40 | 1 | 0 | 0 | 1 |
| Total | 803 | 640 | 467 | 1,910 |

Deaths are concentrated in the first few days. In Table 2.3 with data from the Senegal 2010-11 DHS survey, 212 deaths, 11 percent of the total, occur on the day of birth. The table shows marked heaping at 7 days (column 1, row 07). The number of deaths at 7 days is reported to be 52, which is considerably more than at 6 days or 8 days. It is plausible that some transfers come from 5 days and 9 days, as well, and perhaps even a wider range. The early neonatal period includes days 0-6. If some deaths that actually occur in that interval are incorrectly reported at day 7, then there will be an under-estimate of early neonatal mortality and an over-estimate of late neonatal mortality (in the remainder of the first month).

Table 2.3 also shows pronounced heaping at 12 months (column 2, row 12). Fifty-four deaths are given for that age, far more than at the months just before or after, and it is likely that the transfers are spread over a wider range of months. Deaths that in fact occurred before 12 months, but are reported at 12 months, will lead to an under-estimate of the infant mortality rate.

Table 2.3 includes some violations of the instructions that could have occurred during fieldwork or during data entry. If age at death is reported in days, the maximum value should be 31. Instead, one death is reported at age 36 days, and another at 40 days; both should have been reported at 1 month. If age is reported in months, the maximum value should be 23, but seven deaths reported in months are in the range 24 to 28. Sixty-six deaths in the "years" column are reported at age 1 year. Such deaths should have been reported in months, in the range 12 to 23, but, as will be described below, that level of detail (months in the range 12 to 23) is not actually used in the calculation of mortality rates. There are 11 cases in the "years" column with a value of 10 or greater, even though the maximum possible age at death for a child born in the past 10 years would be 9. Data users have access to b6 but for analytical purposes would use b7.

The second variable describing age at death, b7, or age at death in completed months, is an edited recode of b6. With few exceptions, the conversion from b6 to b7 follows three rules.

Rule 1: If b6 is given in days (first digit 1), then b7 is coded 0 for days 0-29 and 1 for days>=30

Rule 2: If b6 is given in months (first digit 2), then b7=b6-200

Rule 3: If b6 is given in years (first digit 3), then b7=12*(b6-300).

In the Senegal 2010-11 survey, nearly all values of b7 can be calculated by following these rules. For example, the 66 values with b6=301, i.e., age 1 year at death, are converted directly to b7=12 months. The exceptions are due to inconsistencies detected during computer editing, such as the ones identified above in the description of b6. Thus the 11 cases with b6 in the range 310 to 324 are assigned, on the basis of other internal evidence, to values of b7 ranging from 26 to 116 months, and the case with b6=998 is assigned to b7=1. The effect of these deviations from the rules is negligible.

**Table 2.4. Unweighted frequency distribution of b7, imputed months of age at death, in the Senegal 2010-11 DHS Survey, limited to children who were born during the 10 years (120 months) before the mother's interview and who died before the interview**

| b7: Age at death, months | Deaths |
|---|---|
| 0 | 796 |
| 1 | 97 |
| 2 | 71 |
| 3 | 60 |
| 4 | 41 |

**(Continued...)**

**Table 2.4. – *Continued***

| b7: Age at death, months | Deaths |
|---|---|
| 5 | 40 |
| 6 | 47 |
| 7 | 29 |
| 8 | 30 |
| 9 | 35 |
| 10 | 12 |
| 11 | 24 |
| 12 | 114 |
| 13 | 16 |
| 14 | 9 |
| 15 | 11 |
| 16 | 9 |
| 17 | 3 |
| 18 | 40 |
| 19 | 2 |
| 20 | 2 |
| 21 | 2 |
| 22 | 4 |
| 23 | 9 |
| 24 | 189 |
| 25 | 2 |
| 26 | 5 |
| 28 | 1 |
| 35 | 1 |
| 36 | 86 |
| 41 | 1 |
| 48 | 42 |
| 55 | 1 |
| 58 | 1 |
| 60 | 29 |
| 64 | 1 |
| 68 | 1 |
| 69 | 1 |
| 72 | 18 |
| 84 | 16 |
| 96 | 6 |
| 101 | 1 |
| 103 | 1 |
| 107 | 1 |
| 108 | 2 |
| 109 | 1 |
| 116 | 1 |
| Total | 1,911 |

The neonatal interval is officially defined to be the first 28 days, i.e., days 0-27. Defining the first month as 0-29 days brings it into close, but not exact, consistency with the official definition. For the calculation of the neonatal mortality rate, DHS counts the first month (b7=0) as the neonatal interval.

The distribution of b7 in the Senegal 2010-11 survey is given in Table 2.4. Full detail is provided because many users of DHS data are not aware of the characteristics of this distribution. With perfect data, for example from a vital statistics or registration system, the distribution would be smooth except for random fluctuations. The observed distribution in virtually any DHS sample is far from smooth, principally in one specific way—the concentration of responses at multiples of 12 months, because after 23 months, it is intended that the responses will be in years, and then be multiplied by 12.

If the purpose of b7 were to give the best possible estimate of age at death in months, either for individual children or for a distribution, then Rule 3 for converting b6 to b7 would be quite different. It would involve imputing age at death, when given in years, into the midpoint of the year of age, or to a random month within that year of age. Even better, that allocation could have a gradient reflecting the tendency for children to die in the first part of any year of age. As it is, Rule 3 produces a downward bias. If someone used b7 to estimate the mean months of age at death, for example for children dying at ages 2-4 years, as it is coded, the estimate would be nearly half a year too low because of the concentrations at 24 and 36 months.

After 11 months, the concentration of months at multiples of 12 is not nearly as serious a problem as it first may appear to be, because when child mortality rates are constructed, the only information in b7 that is used is the *interval* of months in which it is located, not the numerical value. Those intervals, in months, are 0; 1-2; 3-5; 6-11; 12-23; 24-35; 36-47; and 48-59 Heaping at multiples of 12 months that arises through Rule 3 is not considered to be ambiguous and does not introduce bias into DHS procedures. The only ambiguity that is potentially serious is associated with the heaping at month 12 in the "months" column of Table 2.3, because some of those deaths almost certainly occurred in the interval 6-11 months rather than in the interval 12-23 months.

There is potentially a boundary issue with Rule 3 if the reported age at death, in years, is not correctly understood by the respondent to be *completed* years of age. It is possible that the respondent is giving an estimate of elapsed time since the death, rounded to the nearest year, which would induce a downward bias of half a year of age, on average. That is, if "died at age 3 years" is (incorrectly) interpreted as "died 3 years ago," meaning "died between 2½ and 3½ years ago," then approximately half of such children would have died while age 2 and the other half at age 3 under the (correct) interpretation of age as completed years. There is a similar potential bias in reports of age for living children, as well, if completed years of age is confused with rounded years since birth, but we will not attempt to take it into account.

As mentioned above, the birth histories include two variables to describe the reporting of month and year of birth (b10) and age at death (if the child died, b13). The categories of b10 and b13 are given below, and the distributions for the Senegal 2010-11 DHS survey are given in the form of a cross-tabulation in Table 2.5.

Categories of b10, completeness of birth date

1 month and year - information complete

2 month and age - year imputed

3 year and age - month imputed

4 year and age - year ignored

5 year - age/month imputed

6 age - year/month imputed

7 month - age/year imputed

8 none - all imputed

The coding of b10 specifies whether imputation was required to calculate b3 (the century month of birth). The level of information is “complete,” b10=1, when the calendar year and calendar month of birth for the child are provided in the birth history and, if the child is still alive, the age is also given and is consistent with the date of birth and the date of interview. Otherwise, another code is used for b10 and there will be imputation of month of birth and/or year of birth and/or current age. If month and year are given but age is missing, then age is calculated; this action would not even be considered imputation, and is consolidated with b10=1. If month and year of birth and age are given but are inconsistent, then priority is given to age; year is ignored; b10=4. The codes for b10 encompass all logical possibilities for missing information but some are rare. Generally speaking, the higher the code, the less confidence we would have in the imputed value. The imputation procedures used by DHS have been essentially unchanged since the first phase of the project and are described in the basic project documentation (Croft, nd).

Categories of b13, flag for age at death

0 no flag

1 > interview

2 < breastfeeding

3 < age supplemented

4 < first breastfed

5 < last vaccination

6 outside range

7 imputed, units given

8 imputed, no units

The flag for b13 is constructed in a completely different way, because only one piece of information is obtained about the child’s death—namely, age at death. The variable b13 specifies whether the reported age is consistent with other data in the interview related to the age of the child. If there are no inconsistencies, then b13=0. Most inconsistencies will not result in a change in the age at death because that would require giving priority to the other data. If imputation is done, the imputed value will depend on whether or not the units (days, months, or years, see the discussion of b6) were given. Most surveys conducted after phase 4 do not include questions about age at supplementation of diet and duration of breastfeeding, so codes 2, 3, and 4 are now rarely used.

In most surveys, b10 and b13 are coded 1 and 0, respectively, for most children, and relatively few dates of birth and ages at death are imputed. The woman’s own month and year of birth are sometimes incomplete, in which case DHS will impute v008, but the flag for that type of incompleteness will not be used in this report.

Table 2.5 illustrates the levels of b10 and b13, and their pattern of association, for the Senegal 2010-11 DHS survey. If the child survived, then b13 is not applicable, i.e., is missing. Combinations of b10 and b13 in which b13 is not missing refer to children who died. The inclusion of the breastfeeding flag (b13=2) in the Senegal 2010-11 survey is not typical for recent surveys.

**Table 2.5. Unweighted cross-tabulation of b10 (completeness of birth date) and b13 flag for age at death), in the Senegal 2010-11 DHS Survey, limited to children who were born during the 10 years (120 months) before the mother's interview**

| | b13: Flag for age at death | | | | | | |
|---|---|---|---|---|---|---|---|
| **b10: Completeness of birth date** | **No flag** | **> Inter-view** | **< Breast-feeding** | **Outside range** | **Imputed, no units** | **Child Survived** | **Total** |
| month and year - information complete | 1,111 | 12 | 216 | 49 | 0 | 18,737 | 20,125 |
| month and age - year imputed | 0 | 0 | 0 | 0 | 0 | 1 | 1 |
| year and age - month imputed | 0 | 0 | 0 | 0 | 0 | 1,060 | 1,060 |
| year and age - year ignored | 0 | 0 | 0 | 0 | 0 | 1,682 | 1,682 |
| year - age/month imputed | 433 | 3 | 39 | 22 | 0 | 1 | 498 |
| age - year/month imputed | 0 | 0 | 0 | 0 | 0 | 18 | 18 |
| month - age/year imputed | 3 | 0 | 0 | 0 | 0 | 0 | 3 |
| none - all imputed | 20 | 0 | 0 | 2 | 1 | 4 | 27 |
| Total | 1,567 | 15 | 255 | 73 | 1 | 21,503 | 23,414 |

# 3. Methods

The strategy for this analysis can be summarized with the following steps:

Step 1. Compare the events in the birth history with a set of six criteria that are believed to characterize “correct” data.

Step 2. Adjust the events in the birth history to bring them into compliance with the six criteria, one at a time. The adjustments are accomplished by reweighting the births.

Step 3. Repeat Step 2, iteratively, making all adjustments in succession, until convergence is reached and all criteria are satisfied simultaneously.

Each step of this strategy will produce indices of data quality. The first set of indices, coming from Step 1, measures the deviations between the observed birth histories and each specific criterion. The second set of indices, coming from Step 2, measures the potential impact of the deviations from each specific criterion in terms of the number or percentage of births and under-5 deaths that must be added or moved to satisfy the criteria. The third set of indices, coming from Step 3, measures the potential impact of the deviations from all criteria combined in terms of the number or percentage of births and under-5 deaths. The indices in Steps 2 and 3 take into account the number of cases to which the criteria apply.

The “criteria” should be regarded as plausible reference values. As stated earlier, we can never know what the true or correct values are, and in any case the DHS data are subject to both sampling error and a range of non-sampling errors, some of which are virtually impossible to identify.

## 3.1. Criteria for Identifying Potential Omission

In the implementation of Step 1 of the strategy, six criteria will be employed. Each criterion can be stated as a hypothesis. Other criteria could certainly be added to the list, and in that sense the strategy is potentially much broader than the implementation in this report. Three criteria relate to omission of births or child deaths, and three relate to displacement of dates of births or child deaths.

*Omission Type 1. Sex-specific omission of births.* The sex ratio at birth, or SRB, is defined as 100*(number of male births/number of female births). It is hypothesized that when the SRB deviates from a reference value, it is because of a systematic tendency to omit boys or to omit girls. Based on high-quality international data, it is expected that the sex ratio at birth is 103 boys to 100 girls in sub-Saharan Africa, and 105 boys to 100 girls elsewhere. [PLEASE INSERT FOOTNOTE HERE: The authors acknowledge Valerie Delaunay for information about the sex ratio at birth in Niakhar, Senegal.]

If the SRB is higher than the criterion value, girls will be added with the adjustment procedure. If the SRB is low, boys will be added. The adjustment will increase the (weighted) number of births, and therefore increase fertility rates. It will also increase the (weighted) number of children who died, but will only affect mortality rates to the extent that those rates are sensitive to the sex composition of births.

The SRB is known to vary only slightly around the globe. In a comprehensive study of vital registration data (Visaria, 1967), the SRB varied between 102 and 106. However, multiple studies using vital registration data have documented a lower SRB among African and African-American populations.

Several studies of hospital births in Nigeria (Azeez, Akinboro, and Bakare 2007; Sule and Madugu 2004; Ayeni 1975) have given SRBs closer to those of white populations. James (1984) has discussed the greater potential biases in such data from hospital maternities. We have not considered values of the SRB calculated from DHS data or World Fertility Survey data (given, for example, by Garenne, 2008) for the obvious reason that using such sources would contaminate the assessment. To summarize, we will use slightly different standards or criteria for the SRB in sub-Saharan Africa and elsewhere. The criterion value in sub-Saharan Africa will be 103 and the value elsewhere will be 105.

*Omission Type 2. Sex-specific omission of neonatal deaths.* It is hypothesized that when the sex ratio of neonatal deaths deviates from a criterion value, it is because of a systematic tendency to omit boys or omit girls who experienced a neonatal death. It is expected that the sex ratio of neonatal deaths is 150 boys to 100 girls, i.e., that the proportion who are girls is 0.40 and the proportion who are boys is 0.60.

Either boys or girls will be added with the adjustment procedure, in order to bring the balance to the criterion level. The adjustment will increase the number of births, and therefore will increase the fertility rates.

A recent article by Alkema et al. (2014) describes the pattern of variation in the sex ratio at death for children (but does not separate out the first month). Their interpretation of a lower-than-expected sex ratio at death is that there are excess female deaths. Here, our interpretation would be that some boys who died were omitted. Our interpretation of a higher-than-expected sex ratio at death would be that some girls who died were omitted.

Many studies have documented that in the absence of sex-selective infanticide, male newborns have higher mortality than female newborns, for biological reasons. For example, Teitlebaum (1971) found mortality ratios in this age group of 124 to 151, examining vital registration data in five countries of Northern Europe (Denmark, England and Wales, Norway, Sweden, and Scotland), and Ulizzi and Zonta (2002) documented a mortality ratio of 138 from among 20,000 perinatal deaths in Italy. Thus the ratio of male to female deaths in the neonatal period can be used as a check of data quality. The criterion value, 150, may be somewhat high, but we will only focus on large deviations from that criterion.

*Omission Type 3. Omission of neonatal deaths.* It is hypothesized that births resulting in neonatal deaths tend to be omitted, regardless of whether the child is a boy or a girl, leading to an under-estimate of fertility and of neonatal, infant, and under-5 mortality. The criterion for detecting such omissions is not a fixed number, but is more complex, and will be described separately in Section 3.2.

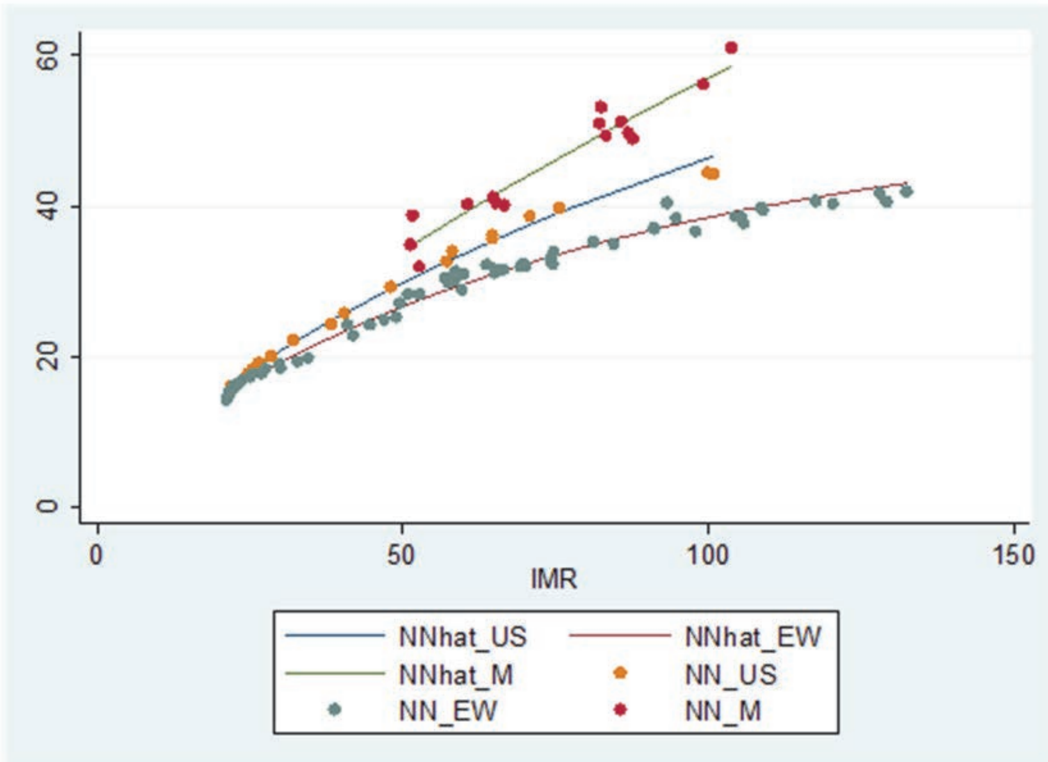
If the number of neonatal deaths appears to be too low, an adjustment will be made to compensate for this deficit. The adjustment will increase the number of births as well as the number of deaths, and therefore will increase both the fertility rates and the mortality rates. No adjustment of the weights is made if the number of neonatal deaths appears to be too high, relative to the expected number.

There may be some reluctance of survey respondents to mention children who died, especially children who were very young when they died. Interviewers may also be reluctant to probe for such events. Few questions in the core DHS questionnaire are asked about children who died before the survey, so the omission of such children will have little effect on the interviewer workload.

### 3.2. Further Detail on Omission of Neonatal Deaths

The criterion for Omission Type 3 is derived from a relationship identified by Hill and Choi (2006). Let NN refer to refer to the neonatal (first month) mortality rate and IMR to the infant (first year) mortality rate. Using historical data from England and Wales, they found that when the IMR is greater than 20 deaths per 1000 births, the ratio NN/IMR fits almost perfectly a linear regression on log(IMR). That is, in symbols, $\frac{NN}{IMR} = b_0 + b_1\log(IMR)$. Hill and Choi did not attempt to motivate this relationship with a model, nor will we, but it is an empirical regularity that appears in other data, including the United States, in the early 20th century, and Matlab, in Bangladesh. Figure 3.1 shows the observed and fitted values of NN and IMR from these three data sets.

**Figure 3.1. Observed and fitted values of the neonatal mortality rate (NN, the vertical axis) and the infant mortality rate (IMR, the horizontal axis), derived from the linear regression of NN/IMR on log(IMR), for IMR>20, using data from England and Wales (EW), the United States (US), and Matlab (M), Bangladesh**




In the original regression of Hill and Choi, the intercept is 1.37 and the coefficient of log(IMR) is -0.214; $R^2$ is 0.98. (These numbers were not in the Hill and Choi paper but were calculated from data kindly provided by Ken Hill. We have applied the model to the U.S. and Matlab data, and we are not aware of anyone else having done so.) In the U.S. data, the intercept is 1.34, and the coefficient of log(IMR) is -0.190; $R^2$ is again 0.98. In the Matlab data, the intercept is 1.30, and the coefficient of log(IMR) is -0.158; $R^2$ is a less impressive 0.55.

In the empirical analysis we explored all three versions and a number of variations with other slopes and intercepts. The initial expectation was that the Matlab coefficients would be most appropriate, because Bangladesh is a developing country, even though its value of $R^2$ is lowest. However, the Matlab coefficients led to implausibly high levels of omission in some countries. The coefficients from the original application (intercept 1.37 and slope -.214) were used because they implied more plausible (and conservative) levels of omission.

Our analysis uses an algebraically equivalent specification of the relationship between neonatal mortality and infant mortality. The regression model given above is rewritten by expressing IMR as NN+PN, where

PN is the postneonatal mortality rate.[10] If we assume that postneonatal mortality is accurately measured,[11] we obtain leverage to estimate an expected value of the neonatal mortality rate. In order to achieve that level, if the observed value is too low, we add children who will then appear both as births and as neonatal deaths. If the observed value of NN is higher than expected, no change is made.

The empirical justification for the model was based on settings in which the IMR was greater than 20 infant deaths per 1,000 live births. For lower values of the IMR, the fit was poorer. There are at least two additional reasons for not applying the model when the IMR is low: reporting tends to be better in such settings, and the small number of deaths in a survey will produce less stable estimates. The model will not be applied to DHS surveys in which the proportion of births in the past 10 years that ended in an infant death is less than 0.03, that is, the approximated IMR is less than 30. The threshold of 30 is arbitrarily set somewhat higher than 20. The check thus omitted 25 surveys in 13 countries: Albania 2008-09; Armenia 2005, 2010; Colombia 1990, 2000, 2005, 2010; Egypt 2008; Honduras 2005-06, 2011-12; Jordan 1997, 2002, 2007, 2012; Kyrgyz Republic 2012; Maldives 2009; Moldova 2005; Peru Continuous Survey, 2004-06, 2007-08, 2010, 2011; Philippines 2003, 2008; Ukraine 2007; Vietnam 2002.

Hill and Choi (2006) also identified an empirical relationship that could allow separate estimation of the early neonatal mortality rate using PN, but we are not using that relationship in this analysis because the double application, first for neonatal mortality and then for early neonatal mortality, leads to adjustments of implausible magnitude.

### 3.3. Criteria for Identifying Potential Displacement

We also hypothesize three types of potential displacement. Displacement is much different from omission because it can be more readily observed as an irregularity in a distribution. The expected distribution is simply a statistical smoothing of the observed distribution.

*Displacement Type 1. Heaping of age at death at 12 months.* In nearly all surveys, children's deaths are disproportionately reported at age 1 year or 12 months. The period of infancy is the first year following the birth, i.e., completed months 0-11. Under the usual demographic interpretation of age, the response "12 months" is 12 *completed* months, and is therefore more than one year after the birth. In DHS surveys there may be considerable heaping of deaths at 12 months. Using the deaths reported in months with b6, we compare the number of deaths in month 12 with one-fifth of the number of deaths in months 10-14.

In the adjustment procedure, deaths within the five-month range of 10 to 14 months, inclusive, are smoothly redistributed across those five months.[12] The effect of the redistribution is to increase the infant mortality rate, to decrease the rate for age 1-4, and to leave the five-year rate for age 0-4 unchanged.

*Displacement Type 2. Heaping of age at death at 7 days.* A similar type of heaping is observed at 7 days. The early neonatal mortality interval is the first 7 days after birth, i.e., completed days 0-6. Early neonatal deaths are defined to occur during the first week. Seven completed days would place the death in the late

---

[10] NN is the proportion of births in the past 10 years that are neonatal deaths, and PN is the proportion of births in the past 10 years that are postneonatal deaths. Because NN is an implicit function of PN, it is necessary to apply an iterative Newton-Raphson procedure to get a fitted value of NN, given the observed value of PN.

[11] If omission *is an issue* for neonatal deaths, then it is a considerable leap of faith to assume that omission *is not an issue* for postneonatal deaths. In practice, the assumption is that omission *is less of an issue* for postneonatal deaths than for neonatal deaths.

[12] The data typically show a declining risk of death from one month of age to the next. The reallocation procedure takes this pattern into account.

neonatal interval. There may be considerable heaping of deaths at 7 days. Using the deaths reported in days with b6, we compare the number of deaths on day 7 with one-fifth of the number of deaths on days 5-9.

In the adjustment procedure, deaths are redistributed in the range of days 5-9.[13] The effect of the redistribution is to increase the early neonatal mortality rate, to decrease the late neonatal rate, and the leave the neonatal rate and all other rates unchanged. The standard list of under-5 rates produced by the DHS does not include the distinction between the early and late neonatal intervals. The criterion is included primarily because heaping at 7 days may suggest more general weaknesses during data collection, rather than because of its potential impact on rates.

*Displacement Type 3. Transfers across the boundary for the health questions*. Types 1 and 2 refer to heaping, in which cases tend to be moved toward a number that represents a rounded response. Type 3 refers to a tendency to move events in one direction. It is hypothesized that births which in fact occurred after the beginning date for the child health questions will tend to be moved to an earlier date, on the other side of the threshold. This kind of displacement is attractive to the interviewer because it substantially reduces the number of health questions to be asked about the child—mainly about a surviving child. As described earlier, the date when the extra questions begin to apply is generally (but not always) January of the fifth calendar year prior to the first month of fieldwork; during data processing some data are discarded so that the reference period for the child is the 59 months prior to the month of interview.[14]

This type of displacement is measured by comparing the number of births reported in the year just before the boundary with the number observed in the first year after the boundary. The adjusted numbers are obtained by smoothing the reported numbers for a four-year interval—the two years before the boundary and the two years after the boundary.

The DHS Woman's Questionnaire has two modules with extensive questions about pregnancy and postnatal care and about child immunizations and health for all births in the five years before the survey. For example, the Malawi 2010 DHS had 71 questions in the pregnancy and postnatal care module and 81 questions in the child immunizations and health module. These are standard questions; sometimes there are additional country-specific questions as well. Many questions only apply conditionally; for example, if the child had diarrhea in the previous two weeks, a set of 13 questions may be asked. Interviewers soon become familiar with the questionnaire and if a mother is not sure of the age of her child but says age 4 or 5, for example, then if the interviewer codes age 5 she can avoid asking the questions in the modules altogether.

### 3.4. Correcting for Omission and Displacement by Altering the Weights

To assess the impact of the six types of errors, we will artificially "correct" the data. This section will describe the procedure for making these adjustments.

If there is evidence that some types of births—for example, births that resulted in a neonatal death—tended to be omitted, a possible strategy to correct the data would be to add some artificial cases to the data file to compensate for the omissions. If there is evidence that some type of birth—for example, a birth that occurred soon after the beginning of the time interval for the health questions—tended to be displaced, a

---

[13] The data typically show a declining risk of death from one day of age to the next. The reallocation procedure takes this pattern into account.

[14] The 60-month interval for the health questions includes the month of interview and the 59 preceding months. The 60-month interval for the calculation of mortality rates omits the month of interview and includes the 60 preceding months. This treatment of the month of interview and the 60th month before the interview can be justified but is a challenge when users of DHS data try to match the numbers in the reports.

possible strategy would be to artificially change the birthdates for some of the births in the interval with an excess of births and thus move them into the interval with a deficit of births.

It is both difficult and inappropriate to add cases or to alter the dates for specific cases. Adding artificial cases with a neonatal death, for example, would require imputing all the b variables, not just the ones for age at death. Shifting birthdates would require some random or arbitrary procedure to select the specific cases to shift, making it difficult for anyone to replicate the results. There would also be a risk of placing two births too close together. An alternative strategy will be used here, namely to manipulate the weight variable, v005, in a manner that simulates the addition or transfer of cases.

DHS surveys have a complex sampling design that typically includes specifying strata (usually all combinations of region and urban-rural residence), selecting clusters within those strata with probability proportional to size, randomly selecting a number of households within each cluster, and then selecting all women age 15-49 for separate interviews that include the birth history questions.

The main reason for using weights is to produce unbiased estimates of population characteristics. If the weights are ignored, then all estimates—proportions, means, regression coefficients, etc., will be biased toward the categories of women who are over-sampled, and away from the categories that are under-sampled. It is standard DHS practice to use weights for most data analyses.[15] The weight variable v005 is normalized by DHS to have a mean of one.[16] Equivalently, the weighted and unweighted total numbers of interviewed women are the same. The weighted and unweighted total numbers of children, who have the same weights as their mothers, may not be exactly the same.

Our strategy will be as follows: If there is evidence of omission for a type of response, then the weight for all cases with that response will be increased by a multiplier to reach the target number of cases. If there is evidence of displacement from one type of response to another, then the weight for all cases with the over-represented response will be decreased by a multiplier, and the weight for all cases with the under-represented response will be increased by a multiplier, calculated to achieve the target number of cases in the two categories, but with no change in the total weight. Adjustments will be made if there is any deviation from a criterion. An alternative strategy would only adjust if the deviation from a criterion were statistically significant, say, or exceeded some threshold level. The difference between the two alternatives—always adjusting or conditionally adjusting—is minor, because if the deviation from the criterion is small, the adjustment will be small.

The remainder of Section 3.4 will describe the process in more detail.

Define a generic variable, $X$, which can have non-missing values $X$=1, 2, or 3 and missing value $X$=".", as in Stata notation. For example, $X$ could be categorized age at death. The missing values would apply to surviving children, with $X$=1 for neonatal deaths, $X$=2 for postneonatal deaths, $X$=3 for later deaths. Say that $W_i$ is the total weight for those cases with $X=i$ (for $i$=1,2,3), and $W_4$ is the total weight (i.e., the sum of the weights) for the cases missing on $X$.

Suppose we have reason to believe that some neonatal deaths have been omitted entirely, both as births and as deaths, and the ratio of neonatal deaths to post-neonatal deaths should be $R$. We want to increase the weights for those children with code $X$=1, with no change to the weights for children with other codes. We

---

[15] For some purposes, data-quality checks do not use weights. The Guide to DHS Statistics states that multivariate analyses conducted by DHS do not use weights, but this is now a minority practice within DHS.

[16] The weights are then multiplied by 1,000,000 and rounded to the nearest integer. The normalization and the factor of 1,000,000 have no effect on calculations in this report.

observe a ratio $r = W_1/W_2$, and need a factor *f* such that when the weights for cases with *X*=1 are multiplied by *f* we will have $R = fW_1/W_2$. It is easily found that $f = R/r$ . If *R*>*r*, which would be consistent with a hypothesis of omission, then *f* will be greater than 1. Thus, the individual weights for all cases with *X*=1 should be multiplied by $f = R/r$. No change would be made to the weights for cases with other values of *X*. The total weight across all cases will have increased because of the adjustment for cases with *X*=1.

Now suppose that the criterion is not specified as a ratio, but as a proportion, for example that the proportion of infant deaths that are neonatal is should be *P*. We observe $p = W_1/(W_1 + W_2)$. We need a factor *f* such that when the weights for cases with X=1 are multiplied by *f* we will have $P = fW_1/(fW_1 + W_2)$. Note that the factor *f* appears in both the numerator and the denominator of the proportion, even though only the cases with code *X*=1 are to be re-weighted. With some algebra it is found that $= \left(\frac{P}{1-P}\right) / \left(\frac{p}{1-p}\right)$. If *P*>$p$, which would be consistent with a hypothesis of omission, then *f* will be greater than 1. Thus, the individual weights for all cases with *X*=1 should be multiplied by this value of $f$. Again, no change should be made to the weights for cases with other values of *X*, but the total weight across all cases will have increased.

The increase in the total weight for the cases with *X*=1 will increase the total sample size. The normalized weights for the entire sample will be changed (that is, will differ from what they were before the weights for cases with *X*=1 were increased). The normalization itself will not be an issue for the calculation of any statistics, except that the additional cases are not "real". That is, the apparent increase in the normalized sample size, by the addition of $W_1(f-1)$ cases, is artificial.

Next suppose that there is evidence that some cases have been transferred from one category to another. For example, it may be reasonable to hypothesize that $p = W_1/(W_1 + W_2)$ is too low, not because of omissions from category 1 but because deaths that should have been reported with *X*=1 were misreported with *X*=2. We need a strategy to increase the weight with *X*=1 and simultaneously reduce the weight with *X*=2 without altering the total weight for the two categories. In this situation, we calculate the expected values of $W_1$ and $W_2$ under a model and refer to them as $E_1$ and $E_2$, respectively. The sum of the *E*s is required to be equal to the sum of the *W*s. Then in the individual-level file, each weight in group *i* (*i*=1,2) will be multiplied by $E_i/W_i$, preserving the total and achieving the desired balance between the two groups. This approach can be expanded to displacement within any number of groups. (This approach can also include omission, described above, by only inflating the weights for the cases in the under-reported category.)

The model to calculate the *E*s in the previous paragraph uses poisson regression. Suppose, for example, that the goal is to adjust for heaping of deaths at 12 months by reallocating the observations at months 10, 11, 12, 13, and 14 in a more regular pattern. The data within that five-month range are collapsed (i.e., the weighted frequencies are summed) to give the observed counts at each of the five values. A poisson regression is fitted to the observed counts. With poisson regression, the fitted values will have the same total as the observed values, providing a simple method to maintain the total count. If the total number of deaths in these five months is less than 50 (an arbitrary cutoff), then they will be allocated equally across the five months—that is, the poisson regression will have no covariates. If the total number of deaths is greater than 50 but the slope increases with age, which is implausible, then they are again allocated equally across the five months. If the total number of deaths is greater than 50 and the slope decreases with age, then age is included as a linear covariate and the best-fitting line (a straight line on a log scale) is used.

### 3.5. Iterative Application of Adjustments

For each criterion, the original data can be adjusted to satisfy that criterion, and the magnitude of the change in the weights can be summarized, as will be described in Section 3.6.

It is also possible to satisfy all the criteria simultaneously. After the data are adjusted to satisfy the first criterion, we can move on to the second adjustment, etc., but each successive adjustment will disrupt the previous ones. In order to satisfy all the criteria simultaneously, it is necessary to cycle through the six adjustments in succession, repeatedly, until convergence is reached—that is, the adjustments made in additional rounds become negligible—and all corrections have been achieved in a mutually compatible manner. It is our experience that this iterative procedure always converges, usually within about six iterations, to any plausible convergence criterion, and the sequence of the adjustments does not matter. There is no requirement or implication that the various adjustments are independent of one another.

The criterion for convergence is based on the mean absolute change in the weights (across the entire sample) from one iteration to the next, beginning with weights that have been initially normalized to a mean of 1. The procedure is terminated when the mean absolute change is less than $10^{-6}$.

Iterative procedures to satisfy multiple criteria have a long history. An early example is provided by Deming (1943, chapter vii), with the adjustment of a two-way table of frequencies to match specified marginal distributions. In that example, the rows and then the columns of the table are adjusted successively and repeatedly until convergence is achieved. Iteratively reweighted least squares (IRLS) is a commonly used computing technique for maximum-likelihood estimation or other kinds of optimization (Basu 2005). Indeed, most estimation models in current use rely on iteration.

As stated above, the adjustment procedure forces the data to match each criterion, even if the original data were already "close"—for example, within one standard error. Otherwise, the iterative application of conditional adjustments would not be expected to converge.

### 3.6. Indicators of the Extent and Impact of Omission and Displacement

The extent of omission and displacement will first be described with indices emerging from Step 1 of the overall strategy that was described at the beginning of Chapter 3. For each of the six criteria and each survey, we will provide either the arithmetic difference or the relative difference between an observed number and an expected number. The indicators will be zero if the observed and expected numbers are equal.

An arithmetic difference is used for the omission criteria, for which the observed and expected numbers are sex ratios or the proportion of infant deaths that are neonatal. These three indices include a factor of 100 and have the form 100*(*O-E*), where *O* is the observed sex ratio or proportion and *E* is the expected or reference value.

A relative difference is used for the displacement criteria. The observed number could be, for example, the number of deaths reported at exactly 12 months. The expected number would then be the number *that would have been* reported at exactly 12 months if the deaths were smoothly distributed over months 10-14, as described earlier. The relative difference, including a factor of 100, has the form 100*(*O-E*)/*E*, where *O* is the observed frequency and *E* is the expected frequency. This form applies also to the number of deaths reported at exactly 7 days.

The indicator of possible displacement across the boundary for the child health questions will be calculated similarly, for the calendar year before the boundary. Say that Year 1 is the calendar year just before the boundary and Year 2 is the first calendar year inside the health window. The observed numbers are O1 and O2, respectively. The expected numbers for both years are assumed to be E1=E2=(O1+O2)/2. The indicator will be defined to be 100*(O1-E1)/E1, which can be shown to be algebraically equivalent to 100(O1-O2)/(O1+O2). A verbal description of the indicator is "the relative deviation of births in the year before the boundary."

Steps 2 and 3 of the overall strategy can also produce indices that measure the extent to which the data depart from the criteria and that will speak more directly to the impact of omission and displacement. Define the individual-level weights in the birth file, before and after adjustment, for birth *i*, as $w_i$ and $w_i$', respectively. Define *A* to be the effective number of children who are added to the file and *D* to be the number of children who have been effectively displaced or moved, by manipulation of the weights. *A* and *D* are calculated from the original and adjusted weights as follows.

*A* is the *net* change in the weights, which will always be an increase, because the omission adjustments only add to the weights: $A = \sum(w_i' - w_i) = W' - W$, where *W* is the sum of the original weights and *W*' is the sum of the adjusted weights.

Define an intermediate quantity, *G*, the sum of the absolute changes in the weights, which is the *gross* change in the weights: $G = \sum|w_i' - w_i|$.

The difference between the gross changes and the net changes, *G-A*, will be the amount of change in the weights that is due to the displacement adjustments. This difference must be divided by 2 because evidence of a displacement appears twice, first as a deficit in one category and then as an excess in another category. Therefore the effective number of displaced births is calculated as $D = [\sum|w_i' - w_i| - \sum(w_i' - w_i)]/2 = \frac{G-A}{2}$. If the calculations are limited to children who died, then *A* and *D* will refer to the numbers of omitted or displaced deaths rather than births. All calculations are limited to children born in the ten years before the survey.

*A* and *D* are interpreted as the weighted numbers of cases added or displaced. If the procedure were repeated with the original weights all set to exactly 1, essentially doing an unweighted analysis, then *A* and *D* could be interpreted as unweighted numbers of cases, even though the adjusted weights would not be 1.

For displacement, the units of distance of the transfer depend on the type of displacement. It can be days (for day of age at death); or months (for month of age at death); or years (a change in the 12-month interval in which the birth occurred).

### 3.7. Calculating Rates with the Adjusted Data

As stated in the introduction, this report will not include a recalculation of fertility and mortality rates. We wish to avoid the risk that recalculated rates would be considered superior to rates that have already been published, or possibly even definitive. Moreover, for many surveys there are probably other kinds of omission and displacement that are more important than the ones included here but that cannot be detected. However, we will briefly describe the linkage between the birth histories and the rates and how the calculation of rates would be affected by changes in the weights.

The fertility and mortality rates relevant to this report, and potentially affected by data quality, are as follows:

*Fertility:* The standard set of seven age-specific rates for five-year intervals of age 15-19 through 45-49; and the total fertility rate (TFR). These rates appear in the main report on every DHS survey.

*Mortality:* The standard set of five age-specific "rates" that appear in the main report on every DHS survey: neonatal; postneonatal; infant (the sum of neonatal and postneonatal); and child (for ages 1-4), and under-5 (for ages 0-4). The infant, child, and under-5 "rates" are actually estimates of the probabilities ${}_1q_0$, ${}_4q_1$, and ${}_5q_0$, in conventional life table notation.

*Time intervals:* DHS reports normally give fertility rates for the three years before the survey, i.e., 0-2 completed years, for the national estimate, urban and rural areas, and major subpopulations. Fertility trends are described with the rates for 0-4, 5-9, 10-14, and 15-19 years before the survey. DHS reports normally present the mortality rates listed above for 0-4 years (i.e., 5 years) before the survey. Mortality rates given for sub-populations are normally for 0-9 years (i.e., 10 years) before the survey. Our estimates of omission and displacement are for the 10 years before the survey. We do not extend farther backward because of the increasing selectivity of surviving respondents and the likely increase in recall error. We do not partition those 10 years into smaller segments, partly because of sample size, especially for the child deaths, and partly because a key issue, displacement at the beginning of the window for the health questions, is located right at the middle of the ten-year interval.

The weight adjustment procedure described in this report applies to the file of births, the "BR" file, by altering the weights attached to the births. The mortality rates for children are always calculated solely from this file, with exactly the same steps either before or after adjustments to the weights.

The fertility rates for women require the events and dates in the "BR" file, but they also require each woman's exposure to risk in each interval of time and age, coming from the "IR" file. In the normal calculation of rates, the weights for women and their children are exactly the same in these two files. However, following the adjustments described in this report, which affect only the "BR" file, the weights for women and their children will not be the same.

It may not be clear how fertility rates could be calculated from the adjusted weights, because the weights in the file of births, which is the source of the numerators of the fertility rates, have been modified, but the weights in the file of women, which is the source of the denominators of the fertility rates, i.e., the exposure component of the rates, are left unchanged. It would appear to be inconsistent to have one set of weights (the original sample weights) for the denominators and another set of weights (the adjusted weights) for the numerators.[17] The following strategy would resolve this inconsistency between the weights for the numerators and denominators. Each birth, which would have a count of 1 with the original calculation, before any weighting or reweighting, would be replaced by the ratio of its adjusted weight to its original weight. The woman's original sampling weights would then be applied to both the numerators and the denominators. In terms of the calculations, this would raise or lower the count of births according to the ratio of the adjusted weight to the original weight, but would not change the denominators or exposure to risk.

Adjustments for birth displacement will, by definition, alter the date of the birth. If the child died, the adjustment may also alter the date of the death that would be inferred from the combination of date of birth and age at death, and if that implied date of death is near the boundary between two successive time intervals, the under-5 probability of death may be altered, going up slightly for one time interval and going down slightly for the other time interval.

[17] Statistical models can be used to calculate the rates—poisson regression for the fertility rates and log probability regression (a type of hazard model) for the mortality rate—but the terminology of numerators and denominators still applies.

# 4. Results of Applying the Criteria for Omission and Displacement: Deviations from the Reference Values

We now turn to the empirical results. For the entire list of 192 surveys, the deviation of each indicator from its hypothetical value will be calculated and the distribution will be displayed in the form of a histogram. Using arbitrary cutoffs, the surveys with the most extreme deviations will be listed. The cutoffs are set simply as values that would identify at most about 10 percent of the surveys and would include the surveys that are clearly in the tails of the distributions. In the lists, the countries are ordered alphabetically according to the country name.

For most indicators there is considerable dispersion around the expected values—much more than could be attributed to sampling variability alone. Even so, large deviations cannot be interpreted as evidence of systematic measurement errors, because some of the deviation may be due to genuinely different values of the characteristic in different countries.

## 4.1. Omission Type 1. Sex-Specific Omission of Births

**Figure 4.1. Distribution of the deviation of the observed sex ratio at birth from the hypothesized value of 105 males (103 in Sub-Saharan Africa) per 100 females, 192 DHS surveys and births in the 10 years before each survey**

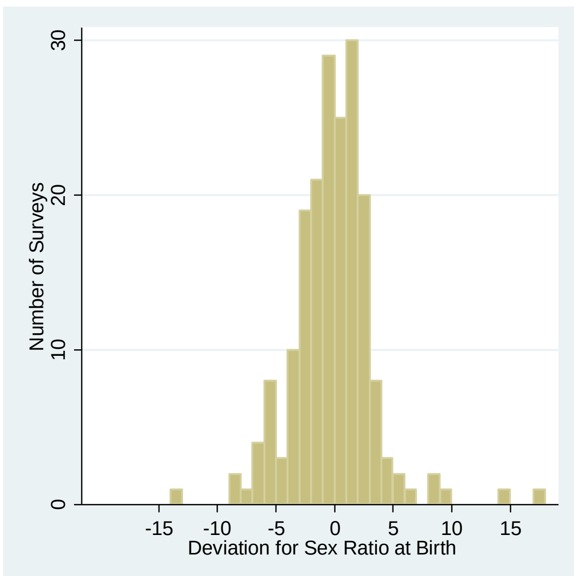



Figure 4.1 shows the distribution of deviations of the sex ratio at birth around the reference values of 103 boys per 100 girls in sub-Saharan Africa and 105 boys per 100 girls elsewhere. The distribution is symmetric, with no evidence of a systematic tendency in DHS surveys to omit girls—or to omit boys.

We have not attempted to test whether the dispersion is greater than would be expected from simply random variation. Such a test would be possible, even though the surveys are of different sizes and have different design effects. However, for reasons given earlier, this kind of a test would not be conclusive. The reference values of 105 and 103 for the sex ratio at birth are themselves approximations to population values that undoubtedly vary across countries and time periods and subgroups.

Tables 4.1a and 4.1b list the surveys with the largest negative and positive deviations from the reference values, with thresholds of -5 points and +5 points, respectively. In a survey with 14,000 births in the past 10 years, approximately the median number for the surveys in this study, and a design effect of 2.0 (so the effective number of births was about 7,000), the half-width of a 95% confidence interval for the sex ratio would be about 5 points. Certainly, a high proportion—perhaps most—of the variation in Figure 4.1 is due to sampling.

Most of the 16 surveys listed in Table 4.1a, which identifies potential omission of boys, are in sub-Saharan Africa (nine surveys) or Latin America and the Caribbean (six surveys). Bangladesh 1996-97 is the only exception. Haiti appears in two surveys and Zambia in three surveys. Because the reference value of the sex ratio at birth was set two points lower for sub-Saharan Africa than for other regions, it is noteworthy that a majority of the surveys on this list of negative deviations are from sub-Saharan Africa. If the reference value for sub-Saharan Africa had been set at a higher value, the preponderance of surveys from that region with large negative deviations would have been even greater. The population of Haiti is primarily of African origin, but the reference value for Haiti was 105. If the reference value had been 103, the Haiti 2000 survey would still be on this list.

Only eight surveys are listed in Table 4.1b, which identifies potential omission of girls. None of them is in sub-Saharan Africa; six are in the North Africa/ West Asia / Europe region. Armenia appears in three surveys. The largest deviation in the list—a surprisingly large amount, 17.0 percent—is for the Armenia 2005 survey. The deviation of 14.3 percent for the Azerbaijan 2006 survey is also very large. These were relatively small surveys. For example, the Armenia 2005 survey included only about 3,000 births in the previous 10 years. The half-width of a 95% confidence interval for the sex ratio would be about 11 points. This type of calculation provides some context but is not conclusive as to the balance between random and systematic variation. Similar kinds of calculations could be made for specific surveys identified elsewhere in this chapter.

In Tables 4.1a and 4.1b, and in the later lists of countries in this chapter, repeat appearances of the same country could indicate either continued problems with data collection or genuine differences in the population from the criterion values. Countries that appear repeatedly, as well as countries for which all the surveys ever conducted appear on the list—even if that is only a single survey—should be investigated further. In the case of Armenia in Table 4.1b, for example, the three surveys are *all* the DHS surveys has conducted in Armenia. Azerbaijan is included only once, but the 2006 survey was the *only* DHS survey in that country. When all the surveys conducted in a country appear in a list, it would appear more likely that the observed deviation is genuine, rather than due to a systematic bias during data collection, but if, say, the same implementing agency was used for all the surveys, it is also possible that the same kinds of measurement errors were repeated.

**Table 4.1a. Surveys with the most extreme negative deviations (fewer boys than expected) from the reference sex ratio at birth, 192 DHS surveys and births in the 10 years before each survey**

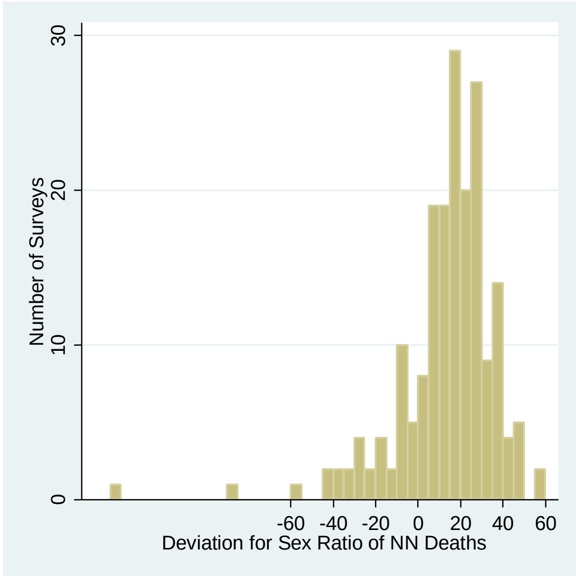
| Survey | Region | Phase of DHS | Deviation for sex ratio at birth |
|---|---|---|---|
| Bangladesh 1996-97 | South and SE Asia | 3 | -5.7 |
| Cameroon 1998 | Middle Africa | 3 | -5.2 |
| Colombia 1990 | Latin America and Carib. | 2 | -8.8 |
| Dominican Republic 1999 | Latin America and Carib. | 4 | -13.4 |
| Gabon 2012 | Middle Africa | 6 | -5.5 |
| Haiti 2000 | Latin America and Carib. | 4 | -8.6 |
| Haiti 2005-06 | Latin America and Carib. | 5 | -6.7 |
| Kenya 1993 | Eastern Africa | 3 | -5.7 |
| Mozambique 2003 | Eastern Africa | 4 | -5.8 |
| Namibia 1992 | Southern Africa | 2 | -6.1 |
| Nicaragua 1998 | Latin America and Carib. | 3 | -5.5 |
| Peru Continuous DHS 2010 | Latin America and Carib. | 6 | -6.3 |
| Uganda 1995 | Eastern Africa | 3 | -7.2 |
| Zambia 1992 | Eastern Africa | 2 | -5.1 |
| Zambia 1996 | Eastern Africa | 3 | -5.8 |
| Zambia 2007 | Eastern Africa | 5 | -6.6 |

**Table 4.1b. Surveys with the most extreme positive deviations (fewer girls than expected) from the reference sex ratio at birth, 192 DHS surveys and births in the 10 years before each survey**

| Survey | Region | Phase of DHS | Deviation for sex ratio at birth |
|---|---|---|---|
| Albania 2008-09 | N Africa / W Asia / Europe | 5 | 5.8 |
| Armenia 2000 | N Africa / W Asia / Europe | 4 | 10.0 |
| Armenia 2005 | N Africa / W Asia / Europe | 5 | 17.0 |
| Armenia 2010 | N Africa / W Asia / Europe | 6 | 8.7 |
| Azerbaijan 2006 | N Africa / W Asia / Europe | 5 | 14.3 |
| Brazil 1991 | Latin America and Carib. | 2 | 5.2 |
| Philippines 2008 | South and SE Asia | 5 | 6.5 |
| Ukraine 2007 | N Africa / W Asia / Europe | 5 | 8.4 |

## 4.2. Omission Type 2. Sex-Specific Omission of Neonatal Deaths

**Figure 4.2. Distribution of the deviation of the observed sex ratio of neonatal deaths from the hypothesized value of 150 males per 100 females, 192 DHS surveys and births in the 10 years before each survey**




The reference value for the sex ratio of neonatal deaths is 150 male deaths per 100 female deaths. Most of the distribution shown in Figure 4.2 is to the right of that reference value. Earlier we suggested that a ratio of 150:100 may be somewhat too high, but, if the 150 were replaced by a smaller number, the displacement to the right would be even greater. There is a great deal of dispersion. The "true" value certainly varies across countries, and when the number of neonatal deaths is relatively small, sampling error can be the main source of dispersion. Tables 4.2a and 4.2b list the surveys with the most extreme deviations, at least 40 points, either below (Table 4.2a) or above (Table 4.2b) the reference value. The Dominican Republic and Vietnam appear in both tables, in different surveys, probably symptomatic of sampling error rather than non-sampling error. The deviations in Tables 4.2a and 4.2b that may warrant further investigation are those for sub-Saharan African countries shown in Table 4.2b, with large positive deviations: Malawi 1992, Mozambique 2003, Niger 1992, Swaziland 2006-07, and Uganda 1995. The extreme positive deviations in Table 4.2b, that is, the surveys with the highest sex ratios for neonatal deaths, suggest greater omission of girls who died in the first month than of boys who died in the first month.

**Table 4.2a. Surveys with the most extreme negative deviations (fewer boys than expected) from the reference sex ratio of neonatal deaths, 192 DHS surveys and births in the 10 years before each survey**

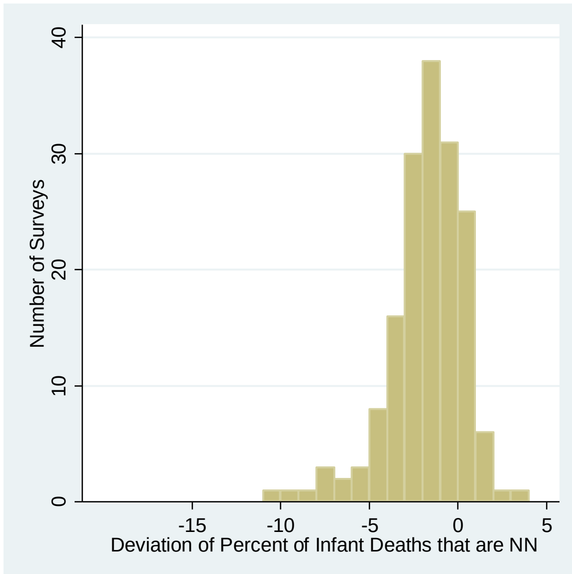
| Survey | Region | Phase of DHS | Deviation for sex ratio of NN deaths |
|---|---|---|---|
| Azerbaijan 2006 | N Africa / W Asia / Europe | 5 | -88.6 |
| Dominican Republic 1991 | Latin America and Carib. | 2 | -43.6 |
| Egypt 2008 | N Africa / W Asia / Europe | 5 | -56.1 |
| Ukraine 2007 | N Africa / W Asia / Europe | 5 | -142.8 |
| Vietnam 1997 | South and SE Asia | 3 | -40.9 |

**Table 4.2b. Surveys with the most extreme positive deviations (fewer girls than expected) from the reference sex ratio of neonatal deaths, 192 DHS surveys and births in the 10 years before each survey**

| Survey | Region | Phase of DHS | Deviation for sex ratio of NN deaths |
|---|---|---|---|
| Bolivia 1994 | Latin America and Carib. | 3 | 41.4 |
| Brazil 1996 | Latin America and Carib. | 3 | 48.6 |
| Colombia 1990 | Latin America and Carib. | 2 | 47.1 |
| Dominican Republic 1999 | Latin America and Carib. | 4 | 58.7 |
| Honduras 2011-12 | Latin America and Carib. | 6 | 45.5 |
| Malawi 1992 | Eastern Africa | 2 | 43.6 |
| Mozambique 2003 | Eastern Africa | 4 | 44.8 |
| Niger 1992 | Western Africa | 2 | 41.0 |
| Swaziland 2006-07 | Southern Africa | 5 | 48.7 |
| Uganda 1995 | Eastern Africa | 3 | 58.1 |
| Vietnam 2002 | South and SE Asia | 4 | 45.8 |

### 4.3. Omission Type 3. Omission of Neonatal Deaths

**Figure 4.3. Distribution of the deviation of the observed ratio of neonatal deaths to infant deaths from the hypothesized ratio (see text), 192 DHS surveys and births in the 10 years before each survey**



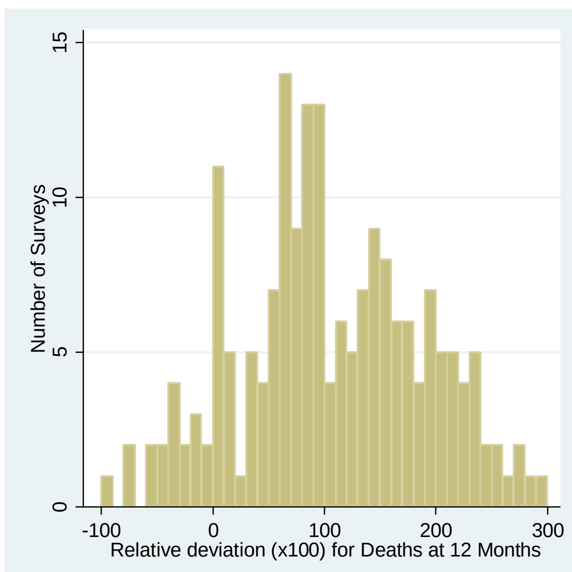

The standard for identifying potential omission of neonatal deaths is the most complex of the three omission criteria. There are reasons to believe that omission would primarily affect deaths in the first day, week, or month of life, but the results shown in Figure 4.3 are concentrated around the reference values, calculated separately for each survey based on the number of infant deaths in the past 10 years, and the proportion of those deaths that are neonatal. Surveys with low infant mortality (that is, in which the approximation to the IMR described in Chapter 3 is less than 30 deaths per 1,000 births) are omitted from this check. A few surveys have small (less than 5 percent) positive deviations, indicating that *more* neonatal deaths are observed than would have been expected. These are surveys with low numbers of infant deaths that almost certainly reflect high variability in small samples. The 11 surveys listed in Table 4.3 are the only ones with a deviation of -5 percent or more (in a negative direction). The Swaziland 2006-07 survey has the largest deviation, -10.2 percent, a likely symptom of this type of omission.

**Table 4.3. Surveys with the most extreme negative deviations (fewer neonatal deaths than expected) from the reference percentage of infant deaths that are neonatal, 192 DHS surveys and births in the 10 years before each survey; 25 surveys with IMR<30 are excluded**

| Survey | Region | Phase of DHS | Deviation of percent of infant deaths that are NN |
|---|---|---|---|
| Brazil 1996 | Latin America and Carib. | 3 | -5.9 |
| Dominican Republic 1999 | Latin America and Carib. | 4 | -9.0 |
| Kazakhstan 1995 | Central Asia | 3 | -7.5 |
| Kenya 1998 | Eastern Africa | 3 | -5.3 |
| Nicaragua 1998 | Latin America and Carib. | 3 | -6.9 |
| Nicaragua 2001 | Latin America and Carib. | 4 | -8.7 |
| Philippines 1993 | South and SE Asia | 3 | -7.5 |
| South Africa 1998 | Southern Africa | 3 | -7.2 |
| Swaziland 2006-07 | Southern Africa | 5 | -10.2 |
| Zambia 2001-02 | Eastern Africa | 4 | -5.6 |
| Zimbabwe 2005-06 | Eastern Africa | 5 | -6.8 |

## 4.4. Displacement Type 1. Heaping of Age at Death at 12 Months

**Figure 4.4. Distribution of the relative deviation of the observed number of deaths at 12 months from the expected number; the expected number is the average for months 10-14, 192 DHS surveys and births in the 10 years before each survey**




The first indicator of displacement is heaping of age at death at 12 months, measured by the relative difference of the number of deaths at 12 months from the average for months 10-14, multiplied by 100. Figure 4.4 gives the distribution of these relative differences across all surveys. For example, "0" in Figure 4.4 means that the number of deaths observed is equal to the number expected; "100" means that the observed number is 100 percent more than expected, i.e., twice as many as expected; "200" means that the observed number is 200 percent more than expected, i.e., three times as many as expected, etc. For a few surveys the indicator is negative; these are surveys with relatively few deaths in the range of 10-14 months, and possibly no deaths reported at exactly 12 months.

Most of the surveys have a substantial level of heaping. Table 4.4 lists surveys for which the indicator is greater than 150. The threshold is set low to include a relatively large number of surveys, nearly a third of the total number, partly to identify countries that appear in two or more surveys: Armenia, Bangladesh, Burkina Faso, Cambodia, Egypt (four times), Ethiopia (three times), Ghana (three times), Guinea, Indonesia (three times), Jordan (four times), Mozambique (three times), Namibia, Niger, Philippines (three times), and Turkey. Of course, countries that have done more DHS surveys are more likely to appear repeatedly, but there are many countries with multiple surveys that do not appear in Table 4.4 at all or appear only once.

Heaping at 12 months in successive surveys is always interpretable as evidence of measurement error, but the error may result from low salience of age and time within the cultural context, rather than from poor interviewing. The largest value of the indicator occurs for the Yemen 1991-92 survey, in which the number of deaths at 12 months was nearly four times as large as expected. Several surveys have an index value of 250 or greater.

Another reason for setting the threshold at 150 is that it is desirable to use the same threshold for both heaping at 12 months and heaping at 7 days, and there is substantially less heaping at 7 days.

**Table 4.4. Surveys with the most extreme heaping on age at death 12 months, 192 DHS surveys and births in the 10 years before each survey**

| Survey | Region | Phase of DHS | Relative deviation (x100) for deaths at 12 months |
|---|---|---|---|
| Armenia 2005 | N Africa / W Asia / Europe | 5 | 251.3 |
| Armenia 2010 | N Africa / W Asia / Europe | 6 | 161.4 |
| Bangladesh 1993-94 | South and SE Asia | 3 | 173.1 |
| Bangladesh 1999-00 | South and SE Asia | 4 | 160.8 |
| Benin 2011-12 | Western Africa | 6 | 164.7 |
| Burkina Faso 1993 | Western Africa | 2 | 222.4 |
| Burkina Faso 2003 | Western Africa | 4 | 164.6 |
| Cambodia 2005 | South and SE Asia | 5 | 167.6 |
| Cambodia 2010 | South and SE Asia | 6 | 184.2 |
| Colombia 1990 | Latin America and Carib. | 2 | 192.8 |
| Comoros 1996 | Eastern Africa | 3 | 194.9 |
| Egypt 1992 | N Africa / W Asia / Europe | 2 | 236.7 |
| Egypt 1995 | N Africa / W Asia / Europe | 3 | 233.9 |
| Egypt 2000 | N Africa / W Asia / Europe | 4 | 258.4 |
| Egypt 2008 | N Africa / W Asia / Europe | 5 | 158.4 |
| Ethiopia 2000 | Eastern Africa | 4 | 152.1 |
| Ethiopia 2005 | Eastern Africa | 5 | 214.8 |
| Ethiopia 2011 | Eastern Africa | 6 | 155.5 |
| Ghana 1993 | Western Africa | 3 | 233.3 |
| Ghana 2003 | Western Africa | 4 | 192.5 |
| Ghana 2008 | Western Africa | 5 | 172.5 |
| Guinea 1999 | Western Africa | 4 | 247.6 |
| Guinea 2012 | Western Africa | 6 | 213.5 |
| India 1998-99 | South and SE Asia | 4 | 181.3 |
| Indonesia 1994 | South and SE Asia | 3 | 209.5 |

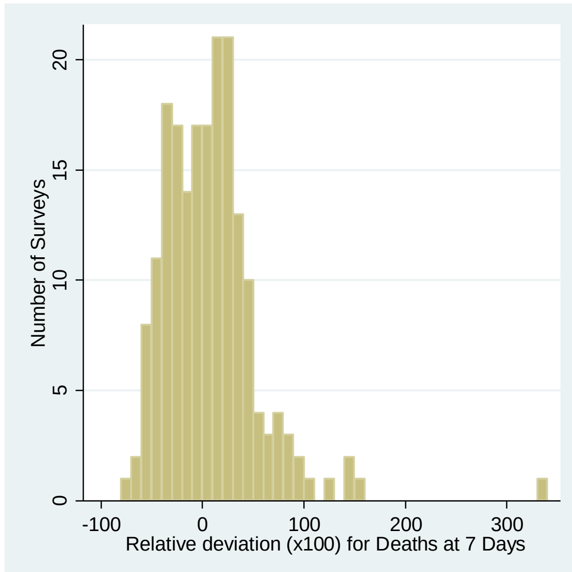
**(Continued...)**

**Table 4.4. – *Continued***

| Survey | Region | Phase of DHS | Relative deviation (x100) for deaths at 12 months |
|---|---|---|---|
| Indonesia 1997 | South and SE Asia | 3 | 186.9 |
| Indonesia 2007 | South and SE Asia | 5 | 212.6 |
| Jordan 1990 | N Africa / W Asia / Europe | 2 | 282.8 |
| Jordan 1997 | N Africa / W Asia / Europe | 3 | 247.1 |
| Jordan 2002 | N Africa / W Asia / Europe | 4 | 231.6 |
| Jordan 2012 | N Africa / W Asia / Europe | 6 | 219.0 |
| Kenya 1993 | Eastern Africa | 3 | 195.9 |
| Mali 2001 | Western Africa | 4 | 204.3 |
| Moldova 2005 | N Africa / W Asia / Europe | 5 | |
| Mozambique 1997 | Eastern Africa | 3 | 174.5 |
| Mozambique 2003 | Eastern Africa | 4 | 197.4 |
| Mozambique 2011 | Eastern Africa | 6 | 200.4 |
| Namibia 1992 | Southern Africa | 2 | 154.0 |
| Namibia 2006-07 | Southern Africa | 5 | 214.3 |
| Niger 2006 | Western Africa | 5 | 276.8 |
| Niger 2012 | Western Africa | 6 | 276.5 |
| Pakistan 2006-07 | South and SE Asia | 5 | 227.8 |
| Peru 1991-92 | Latin America and Carib. | 2 | 151.0 |
| Philippines 1998 | South and SE Asia | 3 | 239.1 |
| Philippines 2003 | South and SE Asia | 4 | 155.2 |
| Philippines 2008 | South and SE Asia | 5 | 170.6 |
| Rwanda 2005 | Eastern Africa | 5 | 154.0 |
| Sao Tome and Principe 2008-09 | Middle Africa | 5 | 265.1 |
| Senegal 1997 | Western Africa | 3 | 181.0 |
| Sierra Leone 2008 | Western Africa | 5 | 208.4 |
| South Africa 1998 | Southern Africa | 3 | 205.7 |
| Swaziland 2006-07 | Southern Africa | 5 | 229.2 |
| Tanzania 1999 | Eastern Africa | 4 | 166.9 |
| Turkey 1993 | N Africa / W Asia / Europe | 3 | 152.2 |
| Turkey 1998 | N Africa / W Asia / Europe | 4 | 190.2 |
| Uganda 2011 | Eastern Africa | 6 | 176.7 |
| Ukraine 2007 | N Africa / W Asia / Europe | 5 | |
| Uzbekistan 1996 | Central Asia | 3 | 173.8 |
| Vietnam 1997 | South and SE Asia | 3 | 225.6 |
| Yemen 1991-92 | N Africa / W Asia / Europe | 2 | 293.5 |
| Zimbabwe 2005-06 | Eastern Africa | 5 | 196.7 |

### 4.5. Displacement Type 2. Heaping of Age at Death at 7 Days

**Figure 4.5. Distribution of the relative deviation of the observed number of deaths at 7 days from the expected number; the expected number is the average for days 5-9, 192 DHS surveys and births in the 10 years before each survey**



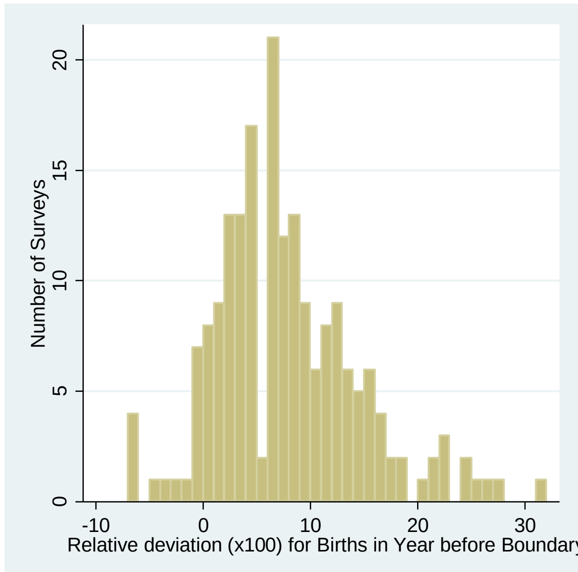

The second indicator of heaping is the concentration of neonatal deaths at 7 days, one day past the boundary for early neonatal deaths (days 0-6). The indicator is calculated as the relative difference of deaths at day 7 from the mean of deaths at days 5-9, multiplied by 100. The histogram in Figure 4.5 shows the distribution of the indicator. It is negative nearly as often as it is positive. There are only two surveys, Armenia 2010 and Ukraine 2007, shown in Table 4.5, for which the indicator exceeds 150, the same threshold that was used for heaping at 12 months. In both surveys the number of deaths during days 5-9 is small. We note that these two surveys were close in both time and region. This type of heaping is less prevalent than was initially suspected.

**Table 4.5. Surveys with the most extreme heaping on age at death 7 days, 192 DHS surveys and births in the 10 years before each survey**

| Survey | Region | Phase of DHS | Relative deviation (x100) for deaths at 7 days |
|---|---|---|---|
| Armenia 2010 | N Africa / W Asia / Europe | 6 | 335.3 |
| Ukraine 2007 | N Africa / W Asia / Europe | 5 | 157.6 |

## 4.6. Displacement Type 3. Transfers across the Boundary for the Health Questions

**Figure 4.6. Distribution of the relative deviation of the observed number of births in the calendar year just before the boundary for the health questions from the expected number, the expected number is the average for the calendar years just before and just after the boundary, 192 DHS surveys and births in the 10 years before each survey**




Of the three types of displacement, transfers across the boundary for the health questions is potentially the most serious. Indicators of such transfers can be calculated in a number of ways. A relatively sophisticated measure, which can be calculated from a logit regression, was described in Pullum (2006). The indicator here is 100 times the relative difference between the observed and expected numbers of births in the calendar year before the boundary. For displacement type 1, the expected value was the mean for months 10-14; for displacement type 2 it was the mean for days 5-9; for displacement type 3 the expected value is the mean for just two years—the years immediately before and after the boundary. As described in Chapter 3, the adjustment procedure employs a poisson regression through four calendar years—the two years before the boundary and the two years after the boundary—but the indicator focuses on just the single year before the boundary and the single year after the boundary.[18]

Figure 4.6 gives the distribution of the 192 surveys across the different levels of displacement. It shows a negative level for a few surveys in which the number of births during the last year before the boundary was *less than* the number in the first year after the boundary. For the great majority of surveys, the number of births before the boundary was at least somewhat above the number after the boundary. Table 4.6 lists the 12 surveys in which the observed number in that year was at least 20 percent greater than expected. The maximum level of displacement was in the Pakistan 1990-91 survey, with an excess of 31.8 percent in the year before the boundary. Apart from that survey and the Cambodia 2000 survey, all of the other 10 surveys were in Western Africa and Eastern Africa. Only one country appears twice—Niger, in both the 2006 and 2012 surveys.

[18] For example, if the first month for the health questions is January 2005, then the year before the boundary is 2004 and the year after the boundary is 2005.

**Table 4.6. Surveys with the most extreme evidence of displacement of births across the boundary for the health questions, 192 DHS surveys and births in the 10 years before each survey**

| Survey | Region | Phase of DHS | Relative deviation (x100) for births in year before boundary |
|---|---|---|---|
| Burkina Faso 1993 | Western Africa | 2 | 21.2 |
| Cambodia 2000 | South and SE Asia | 4 | 21.7 |
| Ghana 1993 | Western Africa | 3 | 22.9 |
| Liberia 2007 | Western Africa | 5 | 24.5 |
| Madagascar 2003-04 | Eastern Africa | 4 | 22.5 |
| Malawi 2000 | Eastern Africa | 4 | 22.4 |
| Mali 2006 | Western Africa | 5 | 24.2 |
| Mozambique 1997 | Eastern Africa | 3 | 20.8 |
| Niger 2006 | Western Africa | 5 | 25.6 |
| Niger 2012 | Western Africa | 6 | 27.4 |
| Pakistan 1990-91 | South and SE Asia | 2 | 31.8 |
| Sierra Leone 2008 | Western Africa | 5 | 26.3 |

# 5. Results of Applying the Criteria for Omission and Displacement: Estimated Numbers of Cases Omitted or Displaced

Chapter 4 described levels and variation in the indicators of omission and displacement. To the extent that the indicators reflect the overall quality of data, including aspects of quality that cannot possibly be measured directly, the information in Chapter 4 would be sufficient. For example, when a measure of heaping on ages ending in 0 or 5, such as Myers' Index, is calculated, a high value is usually interpreted as an indication that age reporting is poor at all ages, with various misstatements, both systematic and random, that do not even show up as heaping. Most analyses of data quality, such as Pullum's (2006), would stop with such measures.

In presenting the strategy for this report, we argued that the indicators themselves provide an incomplete picture of the potential magnitude of omission and displacement, even just of the specific type they purport to describe. For example, several of the criteria, such as heaping of age at death at 12 months or at 7 days, do not actually affect large numbers of births, or even large numbers of deaths. We now describe the impact of the various values of these indicators in terms of the numbers of affected cases, using the implied changes in the weights.

Chapters 4 and 5 complement one another. Both are incomplete in the sense that they use only a limited set of all possible indicators that could be constructed, both are sensitive to the specification of reference values, and both are affected by a layer of randomness or sampling error, but together they give a more complete impression of the magnitude of omission and displacement. Chapter 5 is especially important for distinguishing between the implications for numbers of births and numbers of deaths.

To identify surveys with "high" levels of possible omission or displacement it is necessary to specify thresholds for "high" values that are somewhat arbitrary. The thresholds that will be used are 3 percent, 6 percent, and 10 percent, depending on the number of surveys that would exceed these thresholds. Tables in the Appendix include all the levels for all the surveys.

### 5.1. Estimated Number of Births Omitted by Each Type of Omission

**Figure 5.1. Distribution of 192 DHS surveys according to the estimated number of births omitted because of deviations from each of the three omission criteria (sex ratio at birth, SR_B; sex ratio of neonatal deaths, SR_NN; proportion of infant deaths that are neonatal, NN), expressed as a percentage of the observed number of births during the 10 years before each survey ***

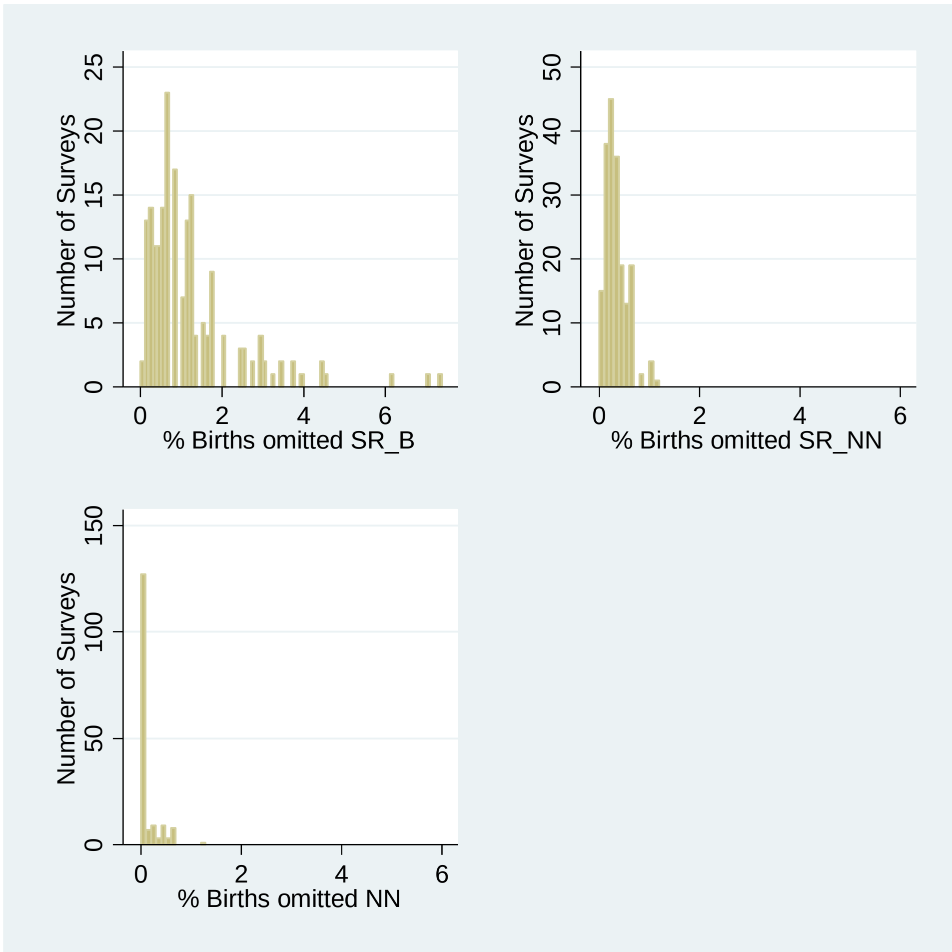


**Table 5.1. Surveys with the highest levels of omission of births implied by the deviation of the sex ratio at birth from the reference value, 192 DHS surveys and births in the 10 years before each survey**

| Survey | Region | Phase of DHS | % Births omitted SR_B |
|---|---|---|---|
| Armenia 2000 | N Africa / W Asia / Europe | 4 | 4.4 |
| Armenia 2005 | N Africa / W Asia / Europe | 5 | 7.3 |
| Armenia 2010 | N Africa / W Asia / Europe | 6 | 3.9 |
| Azerbaijan 2006 | N Africa / W Asia / Europe | 5 | 6.2 |
| Colombia 1990 | Latin America and Carib. | 2 | 4.5 |
| Dominican Republic 1999 | Latin America and Carib. | 4 | 7.0 |
| Haiti 2000 | Latin America and Carib. | 4 | 4.4 |
| Haiti 2005-06 | Latin America and Carib. | 5 | 3.4 |
| Mozambique 2003 | Eastern Africa | 4 | 3.0 |
| Namibia 1992 | Southern Africa | 2 | 3.1 |
| Peru Continuous DHS 2010 | Latin America and Carib. | 6 | 3.2 |
| Uganda 1995 | Eastern Africa | 3 | 3.7 |
| Ukraine 2007 | N Africa / W Asia / Europe | 5 | 3.8 |
| Zambia 2007 | Eastern Africa | 5 | 3.4 |

The three histograms in Figure 5.1 show the number of surveys (the vertical axis) with specified percentages of births estimated to have been omitted because of deviations from the three specified criteria for omission (the horizontal axis). The denominators for these percentages are the observed numbers of births. That is, the horizontal axes of the three figures are the number of cases estimated to have been omitted, divided by the number of observed cases, and multiplied by 100.[19] The three figures are on approximately the same scale, with the horizontal scale extending to 6 percent on all three, to facilitate comparisons. The cutoff for identifying "high" levels of omission in Table 5.1 is arbitrarily set at 3 percent.

The third type of omission, indicated by a low ratio of neonatal to infant deaths, is estimated to be zero for well over half of the surveys, and for that reason the scale of the vertical axis is much different for the third histogram than for the other two histograms in Figure 5.1.

Figure 5.1 shows clearly that variation in the sex ratio at birth is the most important of the three potential indicators of omission of births—although, as stated earlier, variation in the sex ratio at birth can arise for reasons other than omission. In Table 5.1, 14 surveys have implied omission exceeding 3 percent, extending to a maximum level of 7.3 percent in the Armenia 2005 DHS. For the other two indicators of omission, no survey reaches the 3 percent level. The maximum for the other two indicators is about 1 percent.

Two countries appear in repeated surveys in Table 5.1. Haiti appears in two surveys; Armenia appears in three surveys. There is a close correspondence with the surveys and countries listed in Tables 4.1a and 4.1b. Most of the surveys in Table 5.1 had an implied deficit of boys, rather than a deficit of girls, and that is true here as well.

[19] The "percentages" are actually ratios, multiplied by 100, because the denominators are observed numbers of births. In section 5.7 the denominators will be the estimated numbers of births, including those that were omitted.

## 5.2. Estimated Number of Deaths Omitted from Each Type of Omission

**Figure 5.2. Distribution of 192 DHS surveys according to the estimated number of under-5 deaths omitted because of deviations from each of the three omission criteria (sex ratio at birth, SR_B; sex ratio of neonatal deaths, SR_NN; proportion of infant deaths that are neonatal, NN), expressed as a percentage of the observed number of under-5 deaths during the 10 years before each survey**

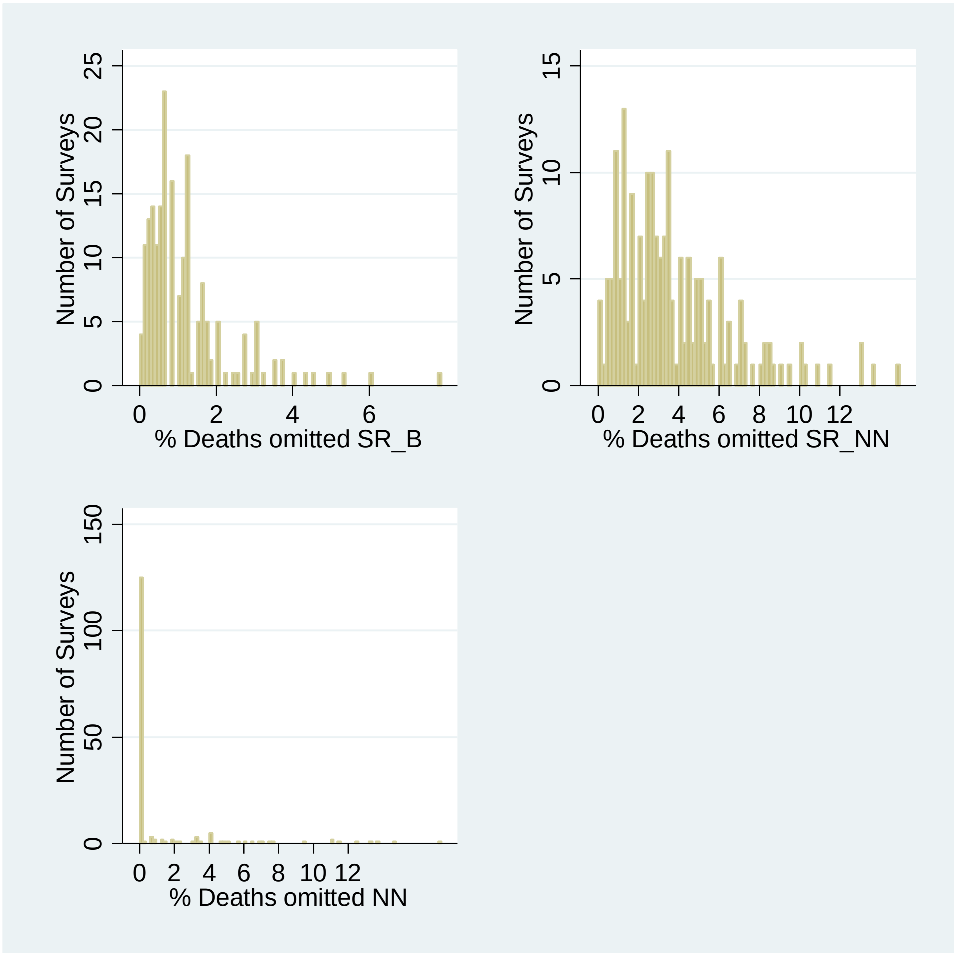


**Table 5.2a. Surveys with the highest levels of omission of under-5 deaths implied by the deviation of the sex ratio at birth from the reference value, 192 DHS surveys and births in the 10 years before each survey**

| Survey | Region | Phase of DHS | % Deaths omitted SR_B |
|---|---|---|---|
| Armenia 2000 | N Africa / W Asia / Europe | 4 | 4.3 |
| Armenia 2005 | N Africa / W Asia / Europe | 5 | 6.0 |
| Armenia 2010 | N Africa / W Asia / Europe | 6 | 4.0 |
| Azerbaijan 2006 | N Africa / W Asia / Europe | 5 | 5.3 |
| Colombia 1990 | Latin America and Carib. | 2 | 4.9 |
| Dominican Republic 1999 | Latin America and Carib. | 4 | 7.8 |
| Gabon 2012 | Middle Africa | 6 | 3.1 |
| Haiti 2000 | Latin America and Carib. | 4 | 4.6 |
| Haiti 2005-06 | Latin America and Carib. | 5 | 3.5 |
| Kenya 1993 | Eastern Africa | 3 | 3.0 |
| Mozambique 2003 | Eastern Africa | 4 | 3.0 |
| Namibia 1992 | Southern Africa | 2 | 3.2 |
| Nicaragua 1998 | Latin America and Carib. | 3 | 3.1 |
| Peru Continuous DHS 2010 | Latin America and Carib. | 6 | 3.6 |
| Uganda 1995 | Eastern Africa | 3 | 3.8 |
| Zambia 1996 | Eastern Africa | 3 | 3.1 |
| Zambia 2007 | Eastern Africa | 5 | 3.7 |

**Table 5.2b. Surveys with the highest levels of omission of under-5 deaths implied by the deviation of the sex ratio of neonatal deaths from the reference value, 192 DHS surveys and births in the 10 years before each survey**

| Survey | Region | Phase of DHS | % Deaths omitted SR_NN |
|---|---|---|---|
| Armenia 2010 | N Africa / W Asia / Europe | 6 | 8.5 |
| Azerbaijan 2006 | N Africa / W Asia / Europe | 5 | 9.5 |
| Bangladesh 1996-97 | South and SE Asia | 3 | 6.0 |
| Bangladesh 2007 | South and SE Asia | 5 | 8.6 |
| Bolivia 1994 | Latin America and Carib. | 3 | 7.0 |
| Brazil 1996 | Latin America and Carib. | 3 | 10.0 |
| Colombia 1990 | Latin America and Carib. | 2 | 10.8 |
| Dominican Republic 1999 | Latin America and Carib. | 4 | 13.2 |
| Egypt 2008 | N Africa / W Asia / Europe | 5 | 6.9 |
| Gabon 2012 | Middle Africa | 6 | 8.4 |
| Guyana 2009 | Latin America and Carib. | 5 | 9.0 |
| Haiti 2005-06 | Latin America and Carib. | 5 | 6.1 |
| Honduras 2011-12 | Latin America and Carib. | 6 | 13.7 |
| India 1998-99 | South and SE Asia | 4 | 6.4 |
| India 2005-06 | South and SE Asia | 5 | 7.2 |
| Indonesia 2002-03 | South and SE Asia | 4 | 6.1 |

**(Continued...)**

**Table 5.2b. – *Continued***

| Survey | Region | Phase of DHS | % Deaths omitted SR_NN |
|---|---|---|---|
| Jordan 1990 | N Africa / W Asia / Europe | 2 | 10.0 |
| Jordan 2002 | N Africa / W Asia / Europe | 4 | 10.3 |
| Jordan 2007 | N Africa / W Asia / Europe | 5 | 7.7 |
| Jordan 2012 | N Africa / W Asia / Europe | 6 | 6.0 |
| Kenya 1993 | Eastern Africa | 3 | 6.0 |
| Kyrgyz Republic 2012 | Central Asia | 6 | 11.5 |
| Lesotho 2004 | Southern Africa | 4 | 6.2 |
| Maldives 2009 | South and SE Asia | 5 | 7.0 |
| Mozambique 2003 | Eastern Africa | 4 | 6.6 |
| Nepal 2001 | South and SE Asia | 4 | 6.4 |
| Nepal 2006 | South and SE Asia | 5 | 8.5 |
| Nepal 2011 | South and SE Asia | 6 | 8.4 |
| Peru Continuous DHS 2004-06 | Latin America and Carib. | 5 | 7.1 |
| Peru Continuous DHS 2007-08 | Latin America and Carib. | 5 | 7.3 |
| Rwanda 1992 | Eastern Africa | 2 | 6.3 |
| Turkey 2003 | N Africa / W Asia / Europe | 4 | 8.1 |
| Uganda 1995 | Eastern Africa | 3 | 7.3 |
| Ukraine 2007 | N Africa / W Asia / Europe | 5 | 15.0 |
| Vietnam 2002 | South and SE Asia | 4 | 13.2 |

**Table 5.2c. Surveys with the highest levels of omission of under-5 deaths implied by the deviation of the proportion of infant deaths that are neonatal from the reference value, 192 DHS surveys and births in the 10 years before each survey**

| Survey | Region | Phase of DHS | % Deaths omitted NN |
|---|---|---|---|
| Brazil 1991 | Latin America and Carib. | 2 | 6.4 |
| Brazil 1996 | Latin America and Carib. | 3 | 9.5 |
| Dominican Republic 1999 | Latin America and Carib. | 4 | 11.2 |
| Kazakhstan 1995 | Central Asia | 3 | 14.6 |
| Kenya 1998 | Eastern Africa | 3 | 7.0 |
| Madagascar 2008-09 | Eastern Africa | 5 | 6.1 |
| Nicaragua 1998 | Latin America and Carib. | 3 | 12.5 |
| Nicaragua 2001 | Latin America and Carib. | 4 | 17.4 |
| Philippines 1993 | South and SE Asia | 3 | 11.5 |
| Philippines 1998 | South and SE Asia | 3 | 7.7 |
| Sao Tome and Principe 2008-09 | Middle Africa | 5 | 7.6 |
| South Africa 1998 | Southern Africa | 3 | 13.4 |
| Swaziland 2006-07 | Southern Africa | 5 | 13.6 |
| Uzbekistan 1996 | Central Asia | 3 | 6.9 |
| Zimbabwe 2005-06 | Eastern Africa | 5 | 11.1 |

We now consider internal evidence of possible omission of under-5 deaths associated with the same three types of potential omission. The first potential indicator of omission, the deviation of the sex ratio at birth from a criterion value, has approximately the same implications for the omission of under-5 deaths as for the omission of births. The other two indicators are defined for deaths, and although they had negligible implications for potential omission of births, they suggest more serious omission of deaths.

Figure 5.2 shows the frequency distributions across all 192 surveys of the potential level of omission of under-5 deaths from these three indicators. As with births, the level of omission for deaths is calculated as a ratio, and is not, strictly speaking, an estimate of the probability of omission. (Estimated probabilities of omission of deaths will be given in Section 5.8.) The horizontal scales for the three parts of Figure 5.2 are different, unfortunately, because of the much different implications for omission of under-5 deaths. The thresholds for identifying "high" values are also different.

Table 5.2a lists 17 surveys for which the potential level of omission of under-5 deaths associated with variation or deviations in the sex ratio at birth is 3 percent or more, with a maximum value of 7.8 percent for the Dominican Republic 1999 survey (a relatively small survey). Again, there are a number of repetitions of the same country in repeated surveys: Armenia, Haiti, and Zambia appear two or three times each. There is a scattering across most DHS geographic regions.

The levels of potential omission of deaths implied by deviations of the sex ratio of neonatal deaths from the criterion value, 150 boys per 100 girls, tend to be higher. The list in Table 5.2b includes 35 surveys with an implied omission of 6 percent or more, reaching a maximum value of 13.7 percent for the Honduras 2011-12 survey. There is a clustering of surveys and countries in South Asia: two surveys in Bangladesh, two in India (that is, both of the India surveys), and three in Nepal. Four surveys from Jordan are included on the list, and two from Peru.

The highest levels of possible omission are reached by the third indicator, the proportion of infant deaths that are neonatal, although only for a few surveys. For most surveys the levels are low. Table 5.2c lists 15 surveys in which the level is more than 5 percent. The list includes eight surveys in which the level is more than 10 percent. The maximum is 17.4 percent in the Nicaragua 2001 DHS. The surveys in Table 5.2c are scattered across all major regions.

We believe that the type of omission described in Table 5.2c is confounded with genuine variation in the proportion of infant deaths that are neonatal, and we consider the results for this indicator to be tentative.

## 5.3. Estimated Number of Births Displaced from Each Type of Displacement

**Figure 5.3. Distribution of 192 DHS surveys according to the estimated number of births displaced because of deviations from each of the three displacement criteria, expressed as a percentage of the observed number of births during the 10 years before each survey**

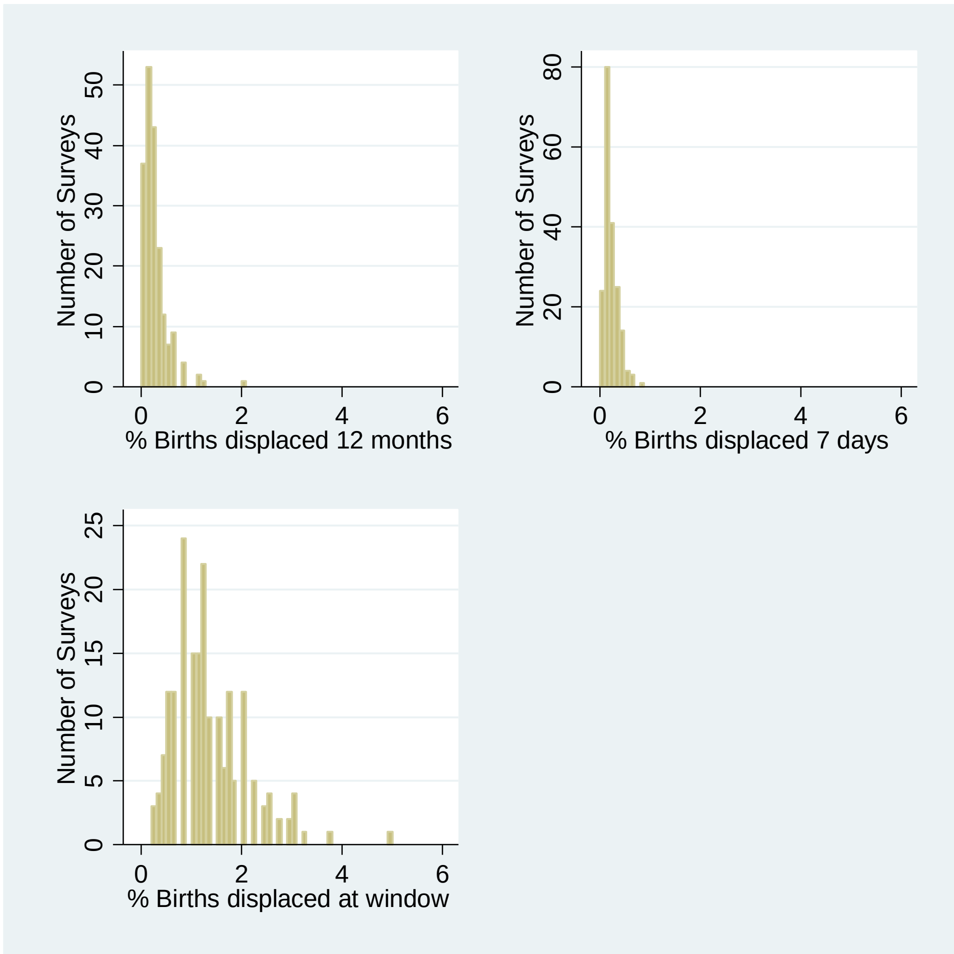



**Table 5.3. Surveys with the highest levels of displacement of births implied by backward transfers across the boundary for the health questions, 192 DHS surveys and births in the 10 years before each survey**

| Survey | Region | Phase of DHS | % Births displaced at window |
|---|---|---|---|
| Cambodia 2000 | South and SE Asia | 4 | 3.8 |
| Ghana 1993 | Western Africa | 3 | 3.1 |
| Liberia 2007 | Western Africa | 5 | 3.0 |
| Mozambique 1997 | Eastern Africa | 3 | 3.0 |
| Niger 2012 | Western Africa | 6 | 3.0 |
| Pakistan 1990-91 | South and SE Asia | 2 | 4.9 |
| Sierra Leone 2008 | Western Africa | 5 | 3.3 |

We now turn to the three types of displacement, one of which is associated with just the displacement of births, and the other two with displacement of deaths as well as births.[20] The distributions of the levels of displacement, first just in terms of births, are shown in the three panels of Figure 5.3. For easier comparison, all three panels have the same scale on the horizontal axis. The thresholds for listing the surveys are all 3 percent. For clarification, the label "% Births displaced 12 months" should be interpreted as 100 times the ratio of (a) the minimum number of births in the past ten years with an age at death in the range of 10 to 14 months that would have to be moved within that range to achieve a smooth distribution within that range, divided by (b) the number of births in the past ten years. The labels "% Births displaced 7 days" and "% Births displaced at window" should be interpreted similarly.

The first type of displacement is only within the range of months of age at death of 10 to 14 months. The second type of displacement is only within the range of days of age at death of 5 to 9 days. The analytical implications of such shifts are negligible for births. The panels in Figure 5.3 for these types of displacement show that the numbers of births that are affected, even within those narrow ranges, are small, at most about 1 percent. The greater impact on numbers of deaths will be described in the next section.

The third type of displacement, across the boundary for the health questions, is of considerably more relevance for births. The displacement affects four calendar years but is mostly from the first year after the boundary to the last year before it. When those years are in different five-year intervals, which is often the case, there are implications for successive five-year fertility rates. Most DHS surveys show at least some displacement.

Table 5.3 lists the surveys—and there are only seven of them—with estimated displacement of 3 percent or more. The largest value is 4.9 percent, for the Pakistan 1990-91 DHS. The other surveys are Ghana 1993, Cambodia 2000, Liberia 2007, Mozambique 1997, Niger 2012, and Sierra Leone 2008. The Niger survey is the only one in the most recent round of surveys, DHS-6.

The small percentages in the third panel of Figure 5.3, and the small number of surveys with more than 3 percent displacement of births, are partly a consequence of the use of 10 years of births as the denominators. The impact on the rates may be more substantial than it first appears, particularly for some comparisons. DHS estimates of trends in fertility are typically based on five-year rates. If, say, 3 percent of the births 0-9 years ago are displaced backward, across the midpoint of the ten-year interval, then the ratio of a fertility rate for 0-4 years ago, to the rate for 5-9 years ago, will be approximately 4*3 percent=12 percent too low, a substantial distortion.[21] Fortunately, only a few surveys appear to have that level of displacement, but a majority of surveys will have some exaggeration of recent declines in fertility, often with a ratio for 0-4 years ago divided by 5-9 years ago that is 4 percent to 8 percent too low (a range of 1 percent to 2 percent in the third panel of Figure 5.3), and occasionally more.

---

[20] We emphasize that the first two types of displacement refer to age at death, for children who died. The deaths are attached to births, but there is no displacement of the birthdates.

[21] Suppose that a proportion $p$ of the births in a ten-year interval have been displaced from 0-4 years ago to 5-9 years ago. Then, in terms of five-year intervals, assuming that about half of the births in the ten-year interval were within the past five years, approximately a proportion $2p$ of the births 0-4 years ago have been shifted to 5-9 years ago. The ratio of births 0-4 years ago to 5-9 years ago will be multiplied by approximately $(1-2p)/(1+2p)$, which is approximately $1-4p$.

## 5.4. Estimated Number of Deaths Displaced from Each Type of Displacement

**Figure 5.4. Distribution of 192 DHS surveys according to the estimated number of under-5 deaths displaced because of deviations from each of the three displacement criteria, expressed as a percentage of the observed number of under-5 deaths during the 10 years before each survey**

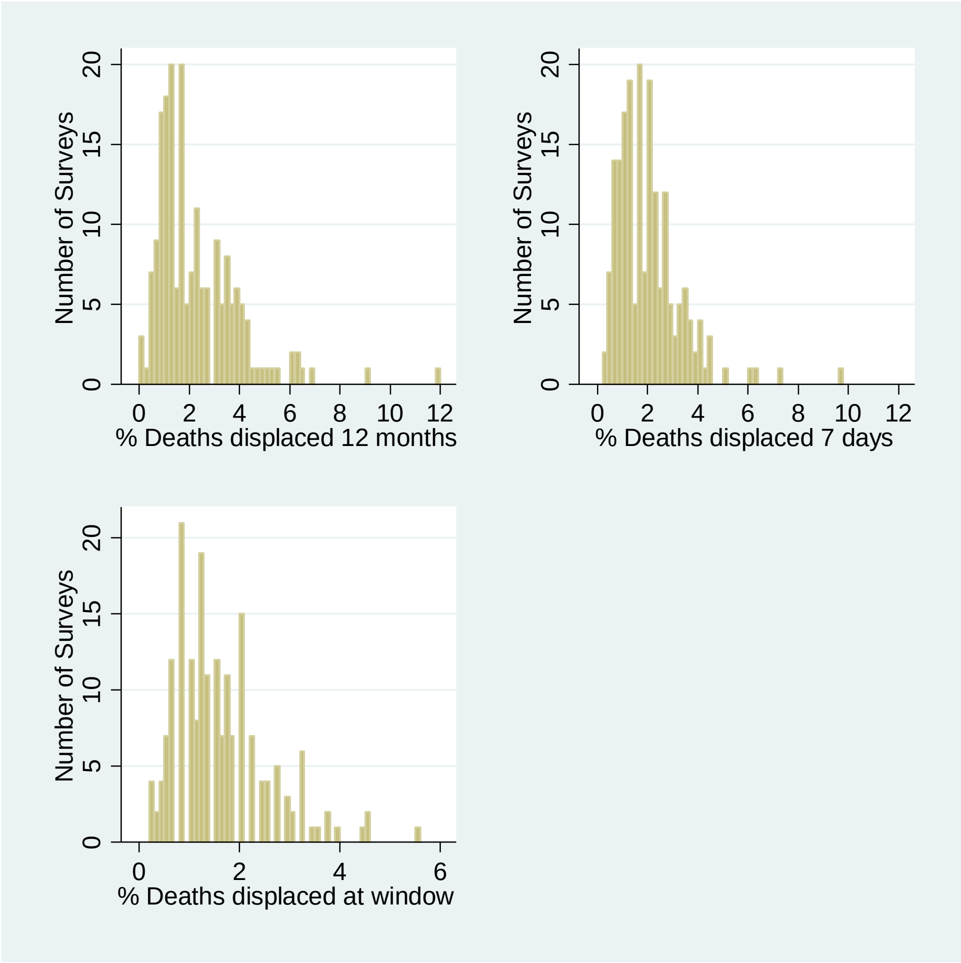


**Table 5.4a. Surveys with the highest levels of displacement of under-5 deaths implied by the heaping of deaths at age 12 months, 192 DHS surveys and births in the 10 years before each survey**

| Survey | Region | Phase of DHS | % Deaths displaced 12 months |
|---|---|---|---|
| Burkina Faso 1993 | Western Africa | 2 | 6.8 |
| Ghana 1993 | Western Africa | 3 | 6.1 |
| Guinea 2012 | Western Africa | 6 | 6.1 |
| Namibia 2006-07 | Southern Africa | 5 | 6.3 |
| Niger 2006 | Western Africa | 5 | 11.8 |
| Niger 2012 | Western Africa | 6 | 9.1 |
| Philippines 1998 | South and SE Asia | 3 | 6.6 |
| Yemen 1991-92 | N Africa / W Asia / Europe | 2 | 6.3 |

**Table 5.4b. Surveys with the highest levels of displacement of under-5 deaths implied by the heaping of deaths at age 7 days, 192 DHS surveys and births in the 10 years before each survey**

| Survey | Region | Phase of DHS | % Deaths displaced 7 days |
|---|---|---|---|
| Armenia 2010 | N Africa / W Asia / Europe | 6 | 7.3 |
| Egypt 1992 | N Africa / W Asia / Europe | 2 | 9.7 |
| Egypt 1995 | N Africa / W Asia / Europe | 3 | 6.2 |
| Egypt 2000 | N Africa / W Asia / Europe | 4 | 6.3 |

**Table 5.4c. Surveys with the highest levels of displacement of under-5 deaths implied by backward transfers across the boundary for the health questions, 192 DHS surveys and births in the 10 years before each survey**

| Survey | Region | Phase of DHS | % Deaths displaced at window |
|---|---|---|---|
| Burkina Faso 1993 | Western Africa | 2 | 3.2 |
| Cambodia 2000 | South and SE Asia | 4 | 4.4 |
| Chad 2004 | Middle Africa | 4 | 3.2 |
| Ghana 1993 | Western Africa | 3 | 3.2 |
| Guinea 2005 | Western Africa | 5 | 3.3 |
| Liberia 2007 | Western Africa | 5 | 4.5 |
| Madagascar 2003-04 | Eastern Africa | 4 | 3.2 |
| Malawi 2000 | Eastern Africa | 4 | 3.1 |
| Mali 2006 | Western Africa | 5 | 3.8 |
| Mozambique 1997 | Eastern Africa | 3 | 3.8 |
| Niger 2006 | Western Africa | 5 | 3.9 |
| Niger 2012 | Western Africa | 6 | 3.6 |
| Nigeria 1990 | Western Africa | 2 | 3.2 |
| Pakistan 1990-91 | South and SE Asia | 2 | 5.6 |
| Peru Continuous DHS 2004-06 | Latin America and Carib. | 5 | 3.4 |
| Rwanda 2000 | Eastern Africa | 4 | 3.0 |
| Sierra Leone 2008 | Western Africa | 5 | 4.6 |

Figure 5.4 shows the distributions of the levels of displacement of deaths. For clarification, the label "% Deaths displaced 12 months" should be interpreted as 100 times the ratio of (a) the minimum number of deaths to children born in the past ten years with an age at death in the range of 10 to 14 months that would have to be moved within that range to achieve a smooth distribution within that range, divided by (b) the number of deaths to children born in the past ten years. The labels "% Deaths displaced 7 days" and "% Deaths displaced at window" should be interpreted similarly.

As would be expected, the criteria for heaping of deaths at 12 months and at 7 days show much more of an effect on the numbers of deaths than was seen for the numbers of births in Section 5.3. The threshold for these two types of displacement is set twice as high as in section 5.3, at 6 percent. In Table 5.4a, the eight surveys with the highest levels of heaping of deaths at 12 months, in terms of the number of deaths that must be reallocated in the 10-14 month range, are Burkina Faso 1993, Ghana 1993, Guinea 2012, two surveys in Niger, 2006 and 2012, Namibia 2006-07, Philippines 1998, and Yemen 1991-92. The two surveys in Niger have the highest levels of all, 11.8 percent and 9.1 percent, respectively.

The four surveys listed in Table 5.4b would require that, of *all* ages at death, at least 6 percent would have to be shifted in order to get a smooth distribution within the interval of 5-9 days. Armenia 2010 is on this list, but the numbers of deaths were relatively small in that survey. Three of the surveys were for Egypt—1992, 1995, and 2000. The highest level of all, 9.7 percent, was in the Egypt 1992 survey. As stated earlier, the only analytical significance of these shifts would be for estimates of early versus late neonatal mortality, which are not normally broken out by DHS, but the relatively high level of heaping could be symptomatic of wider problems with reported age at death.

Table 5.4c gives the list of 17 surveys with more than 3 percent of under-5 deaths displaced. Of these, only four exceed 4 percent and only one exceeds 5 percent. All are in sub-Saharan Africa, except for Cambodia 2000, Pakistan 1990-91, and Peru 2004-06. The highest level in the table is 5.6 percent for Pakistan 1990-91. The only surveys in the same country are Niger 2006 and 2012, both of which also appeared in Table 5.4a. As described for births, 3 percent displacement of deaths, the threshold for Table 5.4c, would imply that the ratio of the under-5 death rate for 0-4 years ago to the rate for 5-9 years ago would be biased downward by up to approximately 4*3 percent=12 percent, which would be considered a serious bias by most standards.

## 5.5. Estimated Number of Births Omitted, Displaced, or Having Incomplete Dates

**Figure 5.5. Distribution of 192 DHS surveys according to the estimated number of births that are omitted, displaced, or are have an incomplete date, expressed as a percentage of the observed number of births during the 10 years before each survey**

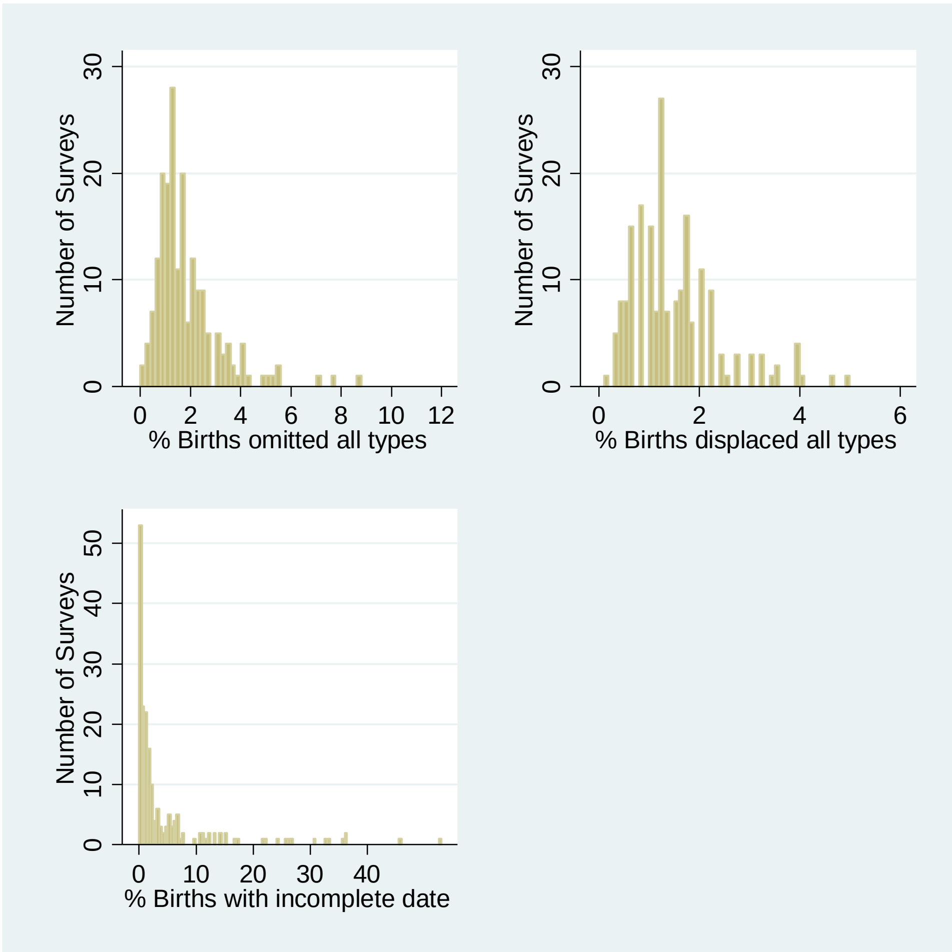



**Table 5.5a. Surveys with the highest levels of omission of births from all three types of omission combined, 192 DHS surveys and births in the 10 years before each survey**

| Survey | Region | Phase of DHS | % Births omitted all types |
|---|---|---|---|
| Armenia 2005 | N Africa / W Asia / Europe | 5 | 7.7 |
| Azerbaijan 2006 | N Africa / W Asia / Europe | 5 | 7.2 |
| Dominican Republic 1999 | Latin America and Carib. | 4 | 8.7 |

**Table 5.5b. Surveys with the highest levels of displacement of births from all three types of displacement combined, 192 DHS surveys and births in the 10 years before each survey**

| Survey | Region | Phase of DHS | % Births displaced all types |
|---|---|---|---|
| Burkina Faso 1993 | Western Africa | 2 | 3.6 |
| Cambodia 2000 | South and SE Asia | 4 | 3.9 |
| Chad 2004 | Middle Africa | 4 | 3.3 |
| Ghana 1993 | Western Africa | 3 | 3.9 |
| Guinea 2005 | Western Africa | 5 | 3.0 |
| Liberia 2007 | Western Africa | 5 | 3.2 |
| Madagascar 2003-04 | Eastern Africa | 4 | 3.0 |
| Malawi 2000 | Eastern Africa | 4 | 3.2 |
| Mali 2006 | Western Africa | 5 | 3.4 |
| Mozambique 1997 | Eastern Africa | 3 | 3.9 |
| Niger 2006 | Western Africa | 5 | 4.7 |
| Niger 2012 | Western Africa | 6 | 4.1 |
| Nigeria 1990 | Western Africa | 2 | 3.1 |
| Pakistan 1990-91 | South and SE Asia | 2 | 5.0 |
| Sierra Leone 2008 | Western Africa | 5 | 3.9 |
| Yemen 1991-92 | N Africa / W Asia / Europe | 2 | 3.5 |

**Table 5.5c. Surveys with the highest levels of incomplete dates of birth, 192 DHS surveys and births in the 10 years before each survey**

| Survey | Region | Phase of DHS | % Births with incomplete date |
|---|---|---|---|
| Benin 1996 | Western Africa | 3 | 36.3 |
| Benin 2001 | Western Africa | 4 | 32.5 |
| Benin 2011-12 | Western Africa | 6 | 26.0 |
| Burkina Faso 1993 | Western Africa | 2 | 21.6 |
| Burkina Faso 1998-99 | Western Africa | 3 | 35.9 |
| Cameroon 1991 | Middle Africa | 2 | 22.1 |
| Comoros 1996 | Eastern Africa | 3 | 25.8 |
| Guinea 1999 | Western Africa | 4 | 52.7 |
| Guinea 2005 | Western Africa | 5 | 36.1 |
| Mozambique 1997 | Eastern Africa | 3 | 26.8 |
| Niger 1992 | Western Africa | 2 | 30.7 |
| Senegal 1992-93 | Western Africa | 2 | 33.3 |
| Togo 1998 | Western Africa | 3 | 24.0 |
| Yemen 1991-92 | N Africa / W Asia / Europe | 2 | 45.9 |

All of the preceding estimates have been specific to each type of omission and displacement. We now turn to estimates after all six criteria have been applied, successively and repeatedly, until convergence. After applying all the criteria repeatedly we can summarize how many births and deaths have potentially been omitted or displaced because of any of the selected mechanisms. We will also include, in Sections 5.5 and 5.6, the percentages of births and deaths with incomplete or flagged information.

The three panels of Figure 5.5 give the potential or implied percentage of births that have been omitted, or displaced, or have incomplete dates of birth, respectively. The denominators are the observed numbers of births in the past 10 years. The panels have different horizontal and vertical scales because the distributions are very different. The surveys with the most extreme values are listed in Tables 5.5a, 5.5b, and 5.5c.

Most of the potential omissions can be traced to variation in the sex ratio at birth, rather than to the other two sources, which primarily affect numbers of deaths. Only three surveys are listed in Table 5.5a, with potential combined omissions of births exceeding 6 percent: Armenia 2005, Azerbaijan 2006, and Dominican Republic 1999. The highest value is 7.1 percent, for the Dominican Republic 1999 DHS.

Most of the displacements of birthdates can be traced to the possible transfers across the boundary for the health questions. Table 5.5b lists 16 surveys with an estimated level exceeding 3 percent. Only three of these surveys exceed 4 percent: Niger 2012 (4.1 percent), Niger 2006 (4.7 percent), and Pakistan 1990-91 (5.0 percent). Niger is the only country included on this list for more than one survey. Except for Cambodia 2000, Pakistan 1990-91, and Yemen 1991-92, all of the surveys in Table 5.5b were conducted in sub-Saharan Africa, nine of them in Western Africa.

Variable b10, which identifies different types of incompleteness of birthdates, was introduced in Chapter 2. Here, it is dichotomized into “complete” (b10=1) and “incomplete” (b10>1). The percentage “incomplete” is calculated for every survey. The third panel of Figure 5.5 shows that the distribution of these percentages is extremely skewed to the right and includes some very high percentages, extending to the largest values of 45.9 percent for Yemen 1991-92 and 52.7 percent for Guinea 1999. Table 5.5c lists the surveys for which the level of incomplete dates of birth was 20 percent or more. There were 14 such surveys, including two in Burkina Faso (1993 and 1998-99), three in Benin (1996, 2001, and 2011-12), and two in Guinea (1999 and 2005). Of the 14 surveys in Table 5.5c, only one—Yemen 1991-92—was outside of sub-Saharan Africa, and 10 of the 13 in sub-Saharan Africa were in Western Africa.

Correlations among these three summary indicators for births were calculated across all surveys. The correlation between potential displacement of birthdates and incomplete birthdates (b10) is 0.33, which is significantly greater than zero at the .0001 level. The correlation between omission and displacement is negative, -.15, which is significant at the .05 level. The third possible correlation, between omission and b10, is not significant.

## 5.6. Estimated Number of Deaths Omitted, Displaced, or Having Flagged Age at Death

**Figure 5.6. Distribution of 192 DHS surveys according to the estimated number of under-5 deaths that are omitted, displaced, or flagged for age at death, expressed as a percentage of the observed number of under-5 deaths during the 10 years before each survey**

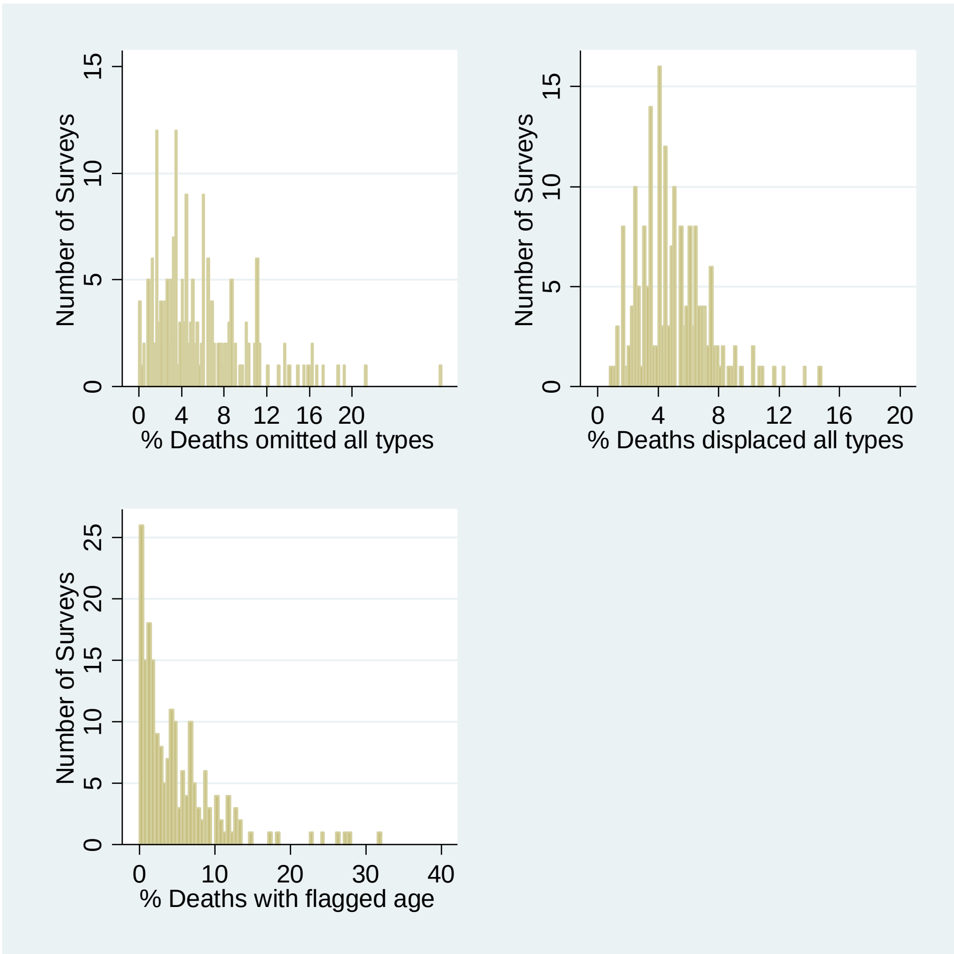


**Table 5.6a. Surveys with the highest levels of omission of under-5 deaths from all three types of omission combined, 192 DHS surveys and births in the 10 years before each survey**

| Survey | Region | Phase of DHS | % Deaths omitted all types |
|---|---|---|---|
| Armenia 2000 | N Africa / W Asia / Europe | 4 | 10.2 |
| Armenia 2005 | N Africa / W Asia / Europe | 5 | 11.1 |
| Armenia 2010 | N Africa / W Asia / Europe | 6 | 10.4 |
| Azerbaijan 2006 | N Africa / W Asia / Europe | 5 | 14 |
| Bangladesh 2007 | South and SE Asia | 5 | 11 |
| Bolivia 1994 | Latin America and Carib. | 3 | 10.2 |
| Brazil 1996 | Latin America and Carib. | 3 | 10.8 |
| Colombia 1990 | Latin America and Carib. | 2 | 21.3 |
| Dominican Republic 1999 | Latin America and Carib. | 4 | 28.2 |
| Gabon 2012 | Middle Africa | 6 | 11.2 |
| Guyana 2009 | Latin America and Carib. | 5 | 11.2 |
| Honduras 2011-12 | Latin America and Carib. | 6 | 17.4 |
| India 2005-06 | South and SE Asia | 5 | 10.2 |
| Jordan 1990 | N Africa / W Asia / Europe | 2 | 13.7 |
| Jordan 2012 | N Africa / W Asia / Europe | 6 | 11.3 |
| Kazakhstan 1995 | Central Asia | 3 | 16.3 |
| Kyrgyz Republic 2012 | Central Asia | 6 | 18.6 |
| Mozambique 2003 | Eastern Africa | 4 | 11 |
| Nepal 2006 | South and SE Asia | 5 | 12 |
| Nepal 2011 | South and SE Asia | 6 | 10.8 |
| Nicaragua 1998 | Latin America and Carib. | 3 | 16.8 |
| Nicaragua 2001 | Latin America and Carib. | 4 | 19.4 |
| Pakistan 2006-07 | South and SE Asia | 5 | 10.3 |
| Philippines 1993 | South and SE Asia | 3 | 13 |
| Philippines 1998 | South and SE Asia | 3 | 11 |
| South Africa 1998 | Southern Africa | 3 | 15.9 |
| Swaziland 2006-07 | Southern Africa | 5 | 15.5 |
| Uganda 1995 | Eastern Africa | 3 | 11.3 |
| Ukraine 2007 | N Africa / W Asia / Europe | 5 | 14.8 |
| Vietnam 2002 | South and SE Asia | 4 | 16.4 |
| Zimbabwe 2005-06 | Eastern Africa | 5 | 13.6 |

**Table 5.6b. Surveys with the highest levels of displacement of under-5 deaths from all three types of displacement combined, 192 DHS surveys and births in the 10 years before each survey**

| Survey | Region | Phase of DHS | % Deaths displaced all types |
|---|---|---|---|
| Burkina Faso 1993 | Western Africa | 2 | 10.6 |
| Egypt 1992 | N Africa / W Asia / Europe | 2 | 13.7 |
| Egypt 1995 | N Africa / W Asia / Europe | 3 | 10.3 |
| Egypt 2000 | N Africa / W Asia / Europe | 4 | 11.6 |
| Ghana 1993 | Western Africa | 3 | 10.8 |
| Niger 2006 | Western Africa | 5 | 14.6 |
| Niger 2012 | Western Africa | 6 | 12.4 |
| Yemen 1991-92 | N Africa / W Asia / Europe | 2 | 10.4 |
| Burkina Faso 1993 | Western Africa | 2 | 10.6 |
| Egypt 1992 | N Africa / W Asia / Europe | 2 | 13.7 |
| Egypt 1995 | N Africa / W Asia / Europe | 3 | 10.3 |
| Egypt 2000 | N Africa / W Asia / Europe | 4 | 11.6 |
| Ghana 1993 | Western Africa | 3 | 10.8 |
| Niger 2006 | Western Africa | 5 | 14.6 |
| Niger 2012 | Western Africa | 6 | 12.4 |
| Yemen 1991-92 | N Africa / W Asia / Europe | 2 | 10.4 |

**Table 5.6c. Surveys with the highest levels of flagged ages at death, 192 DHS surveys and births in the 10 years before each survey**

| Survey | Region | Phase of DHS | % Deaths with flagged age |
|---|---|---|---|
| Armenia 2005 | N Africa / W Asia / Europe | 5 | 31.6 |
| Brazil 1991 | Latin America and Carib. | 2 | 24.0 |
| Colombia 1990 | Latin America and Carib. | 2 | 27.7 |
| Moldova 2005 | N Africa / W Asia / Europe | 5 | 27.1 |
| Nigeria 1999 | Western Africa | 4 | 26.0 |
| Senegal Continuous DHS 2012-13 | Western Africa | 6 | 22.6 |

The three panels of Figure 5.6 provide an overview of the distribution of possible problems with reporting deaths. These distributions reach higher levels than those in Figure 5.5 for births. The three panels refer to potential omission, potential displacement, and the flagging of age at death. This is our first use of flagging of age at death, b10, as an indicator of data quality.

Reference here is to deaths to children born in the 10 years (120 months) prior to the month of interview. The denominators for the three indicators in Figure 5.6 are the observed numbers of deaths. Thus, for example, “the percent of deaths omitted” is actually a ratio, of potentially omitted deaths divided by observed deaths, multiplied by 100, and should not be interpreted as an estimated probability of omission. The probability would be slightly less than the ratio, because the denominator of the probability would be increased to include an estimate of the number of missed deaths. It must be emphasized that the omissions and displacements of deaths include those arising from omissions and displacements of births that resulted in deaths.

Tables 5.6a, 5.6b, and 5.6c list the surveys with the highest levels of potential omission of deaths. Table 5.6a lists 31 surveys with estimated omission levels of 10 percent or more. The surveys are scattered over all regions and phases of DHS, but with only six surveys from sub-Saharan Africa and a conspicuous absence of any surveys in Western Africa. The list includes eight surveys from Latin America and the Caribbean and eight from South and Southeast Asia.

Table 5.6a includes two surveys with potential levels of omission exceeding 20 percent: Colombia 1990 (21.3 percent) and the Dominican Republic 1999 (28.2 percent). We provide the values but do not believe that they can be correctly interpreted as indicating massive omission of under-5 deaths. The Colombia 1990 survey included only 243 under-5 deaths, and the Dominican Republic 1999 survey was particularly small, with only 55 under-5 deaths. These estimates are probably affected seriously by sampling error.

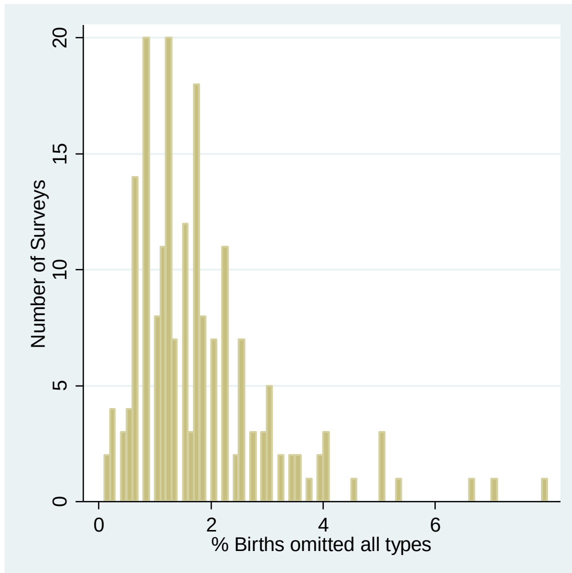
Table 5.6b lists the eight surveys with the highest levels of displacement of under-5 deaths, due to combinations of transfers across the boundary for the health questions plus heaping at age 12 months at death and age 7 days at death. The threshold is again set at 10 percent. The list includes three surveys in Egypt and two in Niger. All the surveys are in Western Africa or in the North Africa / Western Asia / Europe region.

The third panel of Figure 5.6 describes the level of flagging of ages at death, b13, across all surveys. The percentages are lower than was seen earlier for incomplete dates of birth, but are very high for a few surveys. Table 5.6c lists the six surveys with levels of at least 20 percent. The maximum level is 31.6 percent for Armenia 2005. This survey was also highest in Table 5.6b. Colombia 1990 also appears in both Table 5.6b and Table 5.6c; otherwise, the two lists have no surveys in common. The other surveys in Table 5.6c are Brazil 1991, Moldova 2005, Nigeria 1999, and Senegal Continuous DHS 2012-13.

Correlations among the three indicators in Figure 5.6 were calculated using all surveys as units. The flagging of age at death (b13) is not significantly correlated with either omission or displacement. However, omission and displacement have a negative correlation, -.24, which is significant at the .001 level. This relationship is consistent with the negative correlation between omission and displacement of births, described in Section 5.5.

## 5.7. Estimated Probability of Omission of Births

**Figure 5.7. Distribution of 192 DHS surveys according to the estimated probability of omission for births during the 10 years before each survey**




In the remainder of this chapter, omission and displacement of births and deaths will be stated relative to the implied total number of births and deaths in each survey, rather than the observed numbers. They can therefore be interpreted—subject to caveats—as estimated probabilities of omission and displacement. In this section, for example, the estimated number of omitted births is divided by the sum of the observed number of births plus the estimated number of omitted births, and that proportion is multiplied by 100 to get a percentage.

Figure 5.7 shows the distribution of the estimated probability that a birth has been omitted. Table 5.7 lists the 25 surveys with an estimated probability of 3.0 percent or more. Armenia appears three times, for the surveys of 2000, 2005, and 2010. Haiti appears twice, for the surveys in 2000 and 2005-06. Zambia appears twice, for the surveys in 1996 and 2007. Armenia 2005 and Dominican Republic 1999 have the highest values in the table, 6.7 percent and 8.0 percent, respectively.

**Table 5.7. Surveys with the highest estimated probability of omission of births, 192 DHS surveys and births in the 10 years before each survey**

| Survey | Region | Phase of DHS | % Births omitted all types |
|---|---|---|---|
| Armenia 2000 | N Africa / W Asia / Europe | 4 | 4.6 |
| Armenia 2005 | N Africa / W Asia / Europe | 5 | 7.1 |
| Armenia 2010 | N Africa / W Asia / Europe | 6 | 3.9 |
| Azerbaijan 2006 | N Africa / W Asia / Europe | 5 | 6.7 |
| Bangladesh 1996-97 | South and SE Asia | 3 | 3.9 |
| Brazil 1991 | Latin America and Carib. | 2 | 3.0 |
| Cameroon 1998 | Middle Africa | 3 | 3.0 |
| Colombia 1990 | Latin America and Carib. | 2 | 5.1 |
| Dominican Republic 1999 | Latin America and Carib. | 4 | 8.0 |
| Gabon 2012 | Middle Africa | 6 | 3.4 |
| Haiti 2000 | Latin America and Carib. | 4 | 5.0 |
| Haiti 2005-06 | Latin America and Carib. | 5 | 4.1 |
| Kenya 1993 | Eastern Africa | 3 | 3.6 |
| Mozambique 2003 | Eastern Africa | 4 | 5.0 |
| Namibia 1992 | Southern Africa | 2 | 4.0 |
| Nepal 2001 | South and SE Asia | 4 | 3.3 |
| Nicaragua 1998 | Latin America and Carib. | 3 | 3.6 |
| Niger 1992 | Western Africa | 2 | 3.1 |
| Pakistan 2006-07 | South and SE Asia | 5 | 3.4 |
| Peru Continuous DHS 2010 | Latin America and Carib. | 6 | 3.1 |
| Sierra Leone 2008 | Western Africa | 5 | 3.3 |
| Uganda 1995 | Eastern Africa | 3 | 5.3 |
| Ukraine 2007 | N Africa / W Asia / Europe | 5 | 4.0 |
| Zambia 1996 | Eastern Africa | 3 | 3.1 |
| Zambia 2007 | Eastern Africa | 5 | 3.8 |

As with the other lists of surveys in Chapters 4 and 5 of this report, Table 5.7 should be viewed with caution. It is based on criteria for omission that are imperfect and incomplete. To be included on this list, a survey simply needs to show a relatively large deviation from reference values for a combination of sex ratio at birth, sex ratio of neonatal deaths, and the proportion of infant deaths that are in the first month. Alternative criteria and alternative reference values could yield different results.

## 5.8. Estimated Probability of Displacement of Births

**Figure 5.8. Distribution of 192 DHS surveys according to the estimated probability of displacement for births during the 10 years before each survey**

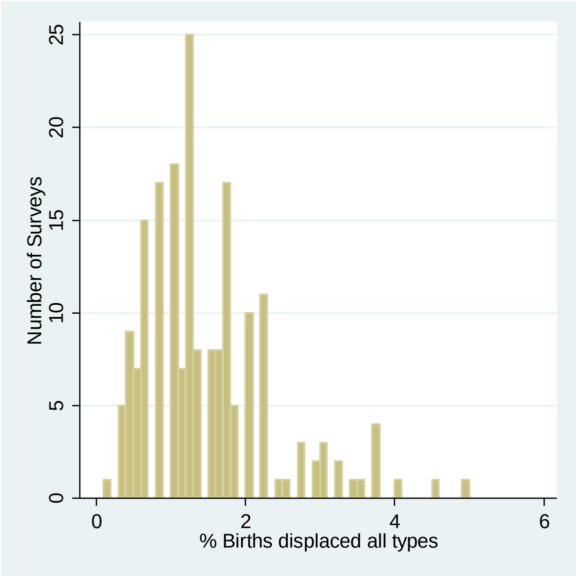


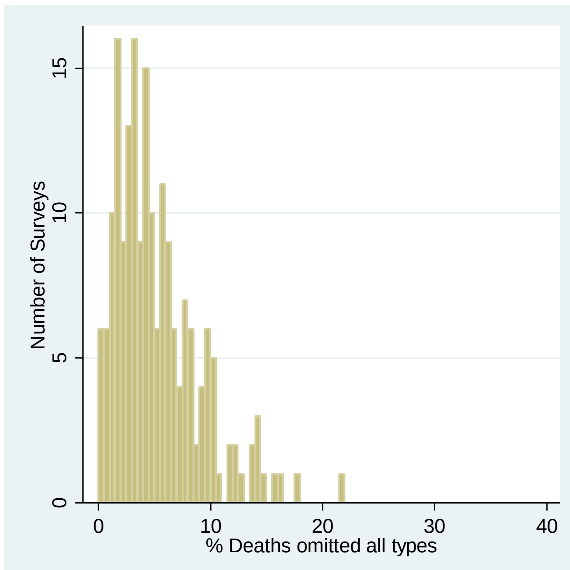

**Table 5.8. Surveys with the highest estimated probability of displacement of births, 192 DHS surveys and births in the 10 years before each survey**

| Survey | Region | Phase of DHS | % Births displaced all types |
|---|---|---|---|
| Burkina Faso 1993 | Western Africa | 2 | 3.5 |
| Cambodia 2000 | South and SE Asia | 4 | 3.8 |
| Chad 2004 | Middle Africa | 4 | 3.2 |
| Ghana 1993 | Western Africa | 3 | 3.8 |
| Liberia 2007 | Western Africa | 5 | 3.1 |
| Malawi 2000 | Eastern Africa | 4 | 3.1 |
| Mali 2006 | Western Africa | 5 | 3.3 |
| Mozambique 1997 | Eastern Africa | 3 | 3.8 |
| Niger 2006 | Western Africa | 5 | 4.6 |
| Niger 2012 | Western Africa | 6 | 4.1 |
| Nigeria 1990 | Western Africa | 2 | 3.0 |
| Pakistan 1990-91 | South and SE Asia | 2 | 4.9 |
| Sierra Leone 2008 | Western Africa | 5 | 3.7 |
| Yemen 1991-92 | N Africa / W Asia / Europe | 2 | 3.4 |

Figure 5.8 shows the estimated probability that a birth has been displaced, and Table 5.8 lists the 14 surveys with estimated probabilities of 3 percent or more. The surveys on this list have been seen before. Except for Cambodia 2000, Pakistan 1990-91, and Yemen 1991-92, all are in sub-Saharan Africa, with eight in Western Africa. The maximum probability of displacement is 4.1 percent for Niger 2006. Almost all of the displacement is by one year, across the boundary for the health questions; only a small amount is associated with displacement toward 12 months or 7 days as ages at death for children who died. As described earlier, a displacement of *p*% across the boundary for the health questions may cause the ratio of the birth rate for 0-4 years before the survey to the birth rate for 5-9 years before the survey to be biased downward by about 4*p*%. Thus for the surveys on this list the ratio may be biased downward by 12 percent or more.

## 5.9. Estimated Probability of Omission of Deaths

**Figure 5.9. Distribution of 192 DHS surveys according to the estimated probability of omission for under-5 deaths during the 10 years before each survey**




The omission of under-5 deaths is related to the omission of births, because if a birth is omitted then the death is also omitted; we assume that it never happens that a birth is reported but the mother claims that the child is still alive, even though in fact the child died—that is, that the death is omitted but the birth is not. The probability that a death is omitted is estimated by dividing the estimated number of omitted deaths by the sum of the observed deaths plus the estimated number of omitted deaths. The only symptoms of omission used in this analysis are a deviation of the sex ratio of neonatal deaths from a criterion value and a deviation of the proportion of infant deaths that are neonatal from a criterion value.

Figure 5.9 suggests that the great majority of surveys have a probability of omission of under-5 deaths that is less than 10 percent. Table 5.9 lists 21 surveys with an estimated probability of 10 percent or more. The only repeats are two surveys from Jordan and two from Nicaragua. There is a scattering across regions, but only five surveys are in sub-Saharan Africa, and none in Western Africa. There is reason to believe that the actual levels of omission of under-5 deaths are highest in sub-Saharan Africa (see Schoumaker, 2014).

**Table 5.9. Surveys with the highest estimated probability of omission of under-5 deaths, 192 DHS surveys and births in the 10 years before each survey**

| Survey | Region | Phase of DHS | % Deaths omitted all types |
|---|---|---|---|
| Armenia 2005 | N Africa / W Asia / Europe | 5 | 10.0 |
| Azerbaijan 2006 | N Africa / W Asia / Europe | 5 | 12.3 |
| Colombia 1990 | Latin America and Carib. | 2 | 17.5 |
| Dominican Republic 1999 | Latin America and Carib. | 4 | 22.0 |
| Gabon 2012 | Middle Africa | 6 | 10.1 |
| Guyana 2009 | Latin America and Carib. | 5 | 10.1 |
| Honduras 2011-12 | Latin America and Carib. | 6 | 14.8 |
| Jordan 1990 | N Africa / W Asia / Europe | 2 | 12.1 |
| Jordan 2012 | N Africa / W Asia / Europe | 6 | 10.1 |
| Kazakhstan 1995 | Central Asia | 3 | 14.1 |
| Kyrgyz Republic 2012 | Central Asia | 6 | 15.7 |
| Nepal 2006 | South and SE Asia | 5 | 10.7 |
| Nicaragua 1998 | Latin America and Carib. | 3 | 14.4 |
| Nicaragua 2001 | Latin America and Carib. | 4 | 16.2 |
| Philippines 1993 | South and SE Asia | 3 | 11.5 |
| South Africa 1998 | Southern Africa | 3 | 13.8 |
| Swaziland 2006-07 | Southern Africa | 5 | 13.5 |
| Uganda 1995 | Eastern Africa | 3 | 10.1 |
| Ukraine 2007 | N Africa / W Asia / Europe | 5 | 12.9 |
| Vietnam 2002 | South and SE Asia | 4 | 14.1 |
| Zimbabwe 2005-06 | Eastern Africa | 5 | 11.9 |

## 5.10. Estimated Probability of Displacement of Deaths

**Figure 5.10. Distribution of 192 DHS surveys according to the estimated probability of displacement for under-5 deaths during the 10 years before each survey**

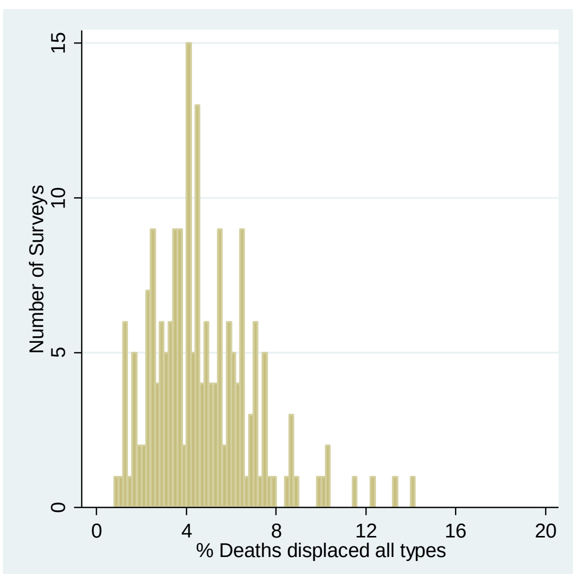


**Table 5.10. Surveys with the highest estimated probability of displacement of under-5 deaths, 192 DHS surveys and births in the 10 years before each survey**

| Survey | Region | Phase of DHS | % Deaths displaced all types |
|---|---|---|---|
| Burkina Faso 1993 | Western Africa | 2 | 10.3 |
| Egypt 1992 | N Africa / W Asia / Europe | 2 | 13.4 |
| Egypt 2000 | N Africa / W Asia / Europe | 4 | 11.5 |
| Ghana 1993 | Western Africa | 3 | 10.3 |
| Niger 2006 | Western Africa | 5 | 14.2 |
| Niger 2012 | Western Africa | 6 | 12.4 |
| Yemen 1991-92 | N Africa / W Asia / Europe | 2 | 10.2 |

The displacement of deaths, like the displacement of births, is easier to identify than omission. Some displacement of deaths, across the boundary for the health questions, arises from the displacement of births. This type of displacement is significant for the identification of trends in fertility and mortality rates. The probability that a birth is displaced is estimated by dividing the number of displaced births by the sum of the observed deaths plus the estimated number of omitted deaths, the same denominator used in Section 5.9.

We also include, as displacement, shifts in age at death that can affect the balance of age categories of under-5 deaths. Heaping at 12 months can affect the assignment of a death between the categories of an infant death (age 0) or a child death (age 1-4). Heaping at 7 days can affect the assignment of the death between the categories of an early neonatal or a late neonatal death.

Figure 5.10 shows the distribution of the estimated probability of displacement of any of these three types. Table 5.10 lists the surveys with an estimated probability of displacement of 10 percent of more. There are 18 such surveys, with a maximum probability of 17.7 percent estimated for the Armenia 2005 survey. Three countries appear in two surveys each: Armenia 2005 and 2010, Egypt 1992 and 2000, and Niger 2006 and 2012. All regions are represented on this list.

# 6. Trends and Regional Variations

Table 2.1 gave the numbers of surveys included in this report in each combination of geographic region—with sub-Saharan Africa divided into four parts—and phase of the DHS project. We will now examine the means of six summary indicators of the quality of the birth histories, in each combination of region and phase, with the goal of identifying trends and regional variations. These means are calculated with equal weight given to every survey, regardless of the number of cases in the survey. In the previous chapters, lists of surveys have been given without taking account of the relative frequencies of the surveys in different time periods and regions. By taking means we will, in effect, adjust for the fact that the numbers of surveys in each phase or region or combination of phase and region have varied widely.

As noted earlier, we are not including any surveys conducted before 1990—that is, in the first phase of DHS, which is known as DHS-I or DHS-1. The subsequent time intervals and labels for the phases will be listed here for reference. The first label is the one that was used at the time of the survey; the second label is the one used at present, consistent with the label "DHS-7" for the current phase, 2014-2018.

| | | |
|---|---|---|
| 1990-1993 | DHS-II | DHS-2 |
| 1994-1998 | DHS-III | DHS-3 |
| 1999-2004 | MEASURE DHS+ | DHS-4 |
| 2005-2009 | MEASURE DHS Phase II | DHS-5 |
| 2010-2013 | MEASURE DHS Phase III | DHS-6 |

There were no surveys in Central Asia in DHS-2 or DHS-5. Those two cells will be empty in the tables in this chapter. The number of surveys in other combinations of geographic region and DHS phase ranges from 1 to 12.

## 6.1. Percentage of Births Having Incomplete Dates

**Table 6.1. Mean percentage of births with incomplete dates (b10), by region and phase of DHS, 192 DHS surveys and births in the 10 years before each survey**

| | Phase of DHS | | | | | |
|---|---|---|---|---|---|---|
| **Region** | **2** | **3** | **4** | **5** | **6** | **All phases** |
| Western Africa | 24.5 | 16.9 | 15.1 | 8.7 | 7.9 | 13.8 |
| Middle Africa | 22.1 | 6.0 | 2.9 | 1.6 | 2.0 | 4.6 |
| Eastern Africa | 7.4 | 9.8 | 2.8 | 1.2 | 0.9 | 4.3 |
| Southern Africa | 5.1 | 2.0 | 1.4 | 0.5 | 0.4 | 1.6 |
| N Africa / W Asia / Europe | 13.7 | 4.9 | 3.4 | 0.8 | 0.0 | 4.2 |
| Central Asia | | 0.3 | 0.0 | | 0.3 | 0.2 |
| South and SE Asia | 5.5 | 1.9 | 1.2 | 3.3 | 0.6 | 2.2 |
| Latin America and Carib. | 0.7 | 1.2 | 0.8 | 0.4 | 0.1 | 0.7 |
| All regions | 10.3 | 6.7 | 4.5 | 2.7 | 2.3 | 4.9 |

Across all 192 surveys, the average percentage of incomplete birth dates is 4.9 percent. Table 6.1 shows a monotonic decline in the percentage over time, from 10.3 percent in DHS-2 to 2.3 percent in DHS-6. Within regions, the decline has not been perfectly monotonic, but in all regions the percentage is much lower in DHS-6 than in DHS-2. The improvement has been greatest in the North Africa / West Asia / Europe region.

In every phase of DHS the percentage of incomplete dates has been highest in Western Africa, averaging 13.8 percent across all the surveys conducted there. It has been about 4 percent in Middle Africa, in Eastern Africa, and in North Africa / West Asia / Europe. Central Asia (0.2 percent) and Latin America (0.7 percent) have consistently had negligible levels of incomplete dates. By DHS-6, the level was below 1.0 percent in all regions except Western Africa and Middle Africa. Thus by DHS-6, the role of computer imputation of birthdates was largely eliminated in all regions except Western Africa. It must be noted, of course, that birthdates can be complete without being accurate.

### 6.2. Percentage of Deaths Having Flagged Age

**Table 6.2. Mean percentage of deaths with flagged age at death (b13), by region and phase of DHS, 192 DHS surveys and births in the 10 years before each survey**

| | Phase of DHS | | | | | |
|---|---|---|---|---|---|---|
| Region | 2 | 3 | 4 | 5 | 6 | All phases |
| Western Africa | 2.9 | 2.7 | 10.6 | 8.6 | 7.9 | 6.9 |
| Middle Africa | 1.3 | 4.7 | 4.0 | 6.3 | 2.9 | 4.2 |
| Eastern Africa | 2.5 | 3.8 | 7.1 | 6.9 | 3.0 | 5.0 |
| Southern Africa | 2.8 | 12.6 | 6.8 | 7.9 | 12.0 | 8.1 |
| N Africa / W Asia / Europe | 2.8 | 1.2 | 3.3 | 10.5 | 0.1 | 5.2 |
| Central Asia | | 2.4 | 3.6 | | 6.5 | 3.4 |
| South and SE Asia | 2.6 | 1.2 | 2.9 | 4.0 | 4.0 | 2.9 |
| Latin America and Carib. | 10.5 | 3.5 | 3.2 | 2.6 | 2.3 | 4.1 |
| All regions | 4.3 | 3.1 | 5.7 | 6.6 | 4.3 | 4.9 |

Ages at death are flagged if the age given by the respondent is inconsistent with other ages or durations in response to the child health questions, for example those questions that elicit responses about immunizations or breastfeeding. Across all 192 surveys, the mean percentage of deaths with flagged age is 4.9 percent, coincidentally the same as the mean percentage of births with incomplete deaths. However, the pattern of such flagging across regions and phases of DHS, shown in Table 6.2, is not as easily summarized as the pattern of incomplete dates of birth.

Overall, the level of flagging in DHS-6, 4.3 percent, is the same as the level in DHS-2. That is, there has been no net reduction over time. Flagging was somewhat more common in DHS-4 (5.7 percent) and DHS-5 (6.6 percent) than in DHS-2, DHS-3, and DHS-6. Latin America is the only region that has shown a monotonic decline in the level of flagging across the full sequence from DHS-2 to DHS-6. All other regions, except North Africa / West Asia / Europe, have a higher level of flagging in DHS-6 than in DHS-2.

In DHS-6, North Africa / West Asia / Europe is the only region to have achieved a negligible level of flagging. By contrast, Southern Africa and Western Africa show the highest levels of flagging in DHS-6, 12.0 percent and 7.9 percent, respectively.

## 6.3. Estimated Probability of Omission of Births

**Table 6.3. Mean estimated probability of omission of a birth, by region and phase of DHS, 192 DHS surveys and births in the 10 years before each survey**

| | Phase of DHS | | | | | |
|---|---|---|---|---|---|---|
| Region | 2 | 3 | 4 | 5 | 6 | All phases |
| Western Africa | 1.9 | 1.2 | 1.2 | 1.6 | 1.2 | 1.4 |
| Middle Africa | 1.4 | 1.8 | 1.1 | 1.3 | 2.2 | 1.6 |
| Eastern Africa | 2.3 | 2.6 | 1.8 | 1.9 | 1.4 | 2.0 |
| Southern Africa | 4.0 | 1.7 | 1.8 | 1.7 | 1.3 | 2.0 |
| N Africa / W Asia / Europe | 0.7 | 1.3 | 1.8 | 2.8 | 2.4 | 1.9 |
| Central Asia | | 1.5 | 1.1 | | 2.5 | 1.6 |
| South and SE Asia | 1.8 | 1.8 | 1.7 | 2.3 | 1.0 | 1.8 |
| Latin America and Carib. | 2.4 | 1.5 | 2.5 | 1.5 | 1.5 | 1.8 |
| All regions | 2.0 | 1.7 | 1.7 | 2.0 | 1.5 | 1.8 |

Sections 6.3 to 6.6 will use the survey-specific estimates of the probabilities of estimation and displacement that correspond with Sections 5.7 to 5.10, respectively. The probabilities are expressed in percentage format—that is, they have been multiplied by 100. It must be repeated that these estimated probabilities are based on the specified criteria and reference values and are limited to that framework. As we have stated repeatedly, omission is particularly difficult to identify.

Table 6.3 suggests that the average probability of omission of births has been 1.8 percent, with relatively little variation across DHS phases and regions. A comparison of DHS-6 with DHS-2 suggests a net decline over time, from 2.0 percent to 1.5 percent, although DHS-5 had an intervening increase to 2.0 percent. Most regions show a net decline, especially Southern Africa, which declined monotonically from 4.0 percent (the highest level in Table 6.3) in DHS-2 to 1.3 percent in DHS-6. Two of the eight regions show an increase from DHS-2 to DHS-6: Middle Africa (from 1.4 percent to 2.2 percent) and North Africa / West Asia / Europe (from 0.7 percent to 2.4 percent). Central Asia did not have any surveys in DHS-2, but shows an increase from DHS-3 to DHS-6 (from 1.5 percent to 2.5 percent).

These mean values are low, and suggest that omission of births is not a serious problem with DHS surveys. Two percent omission, or under-estimation of fertility, is within the range of sampling error for virtually all estimates of fertility rates. However, two things must be kept in mind: first, 2 percent is a mean and several specific surveys have much higher levels of omission; second, the strategy to identify omission has certainly not been complete.

## 6.4. Estimated Probability of Displacement of Births

**Table 6.4. Mean estimated probability of displacement of a birth, by region and phase of DHS, 192 DHS surveys and births in the 10 years before each survey**

| | Phase of DHS | | | | | |
|---|---|---|---|---|---|---|
| **Region** | **2** | **3** | **4** | **5** | **6** | **All phases** |
| Western Africa | 2.4 | 1.9 | 1.5 | 3.0 | 2.0 | 2.1 |
| Middle Africa | 1.3 | 1.5 | 1.9 | 1.4 | 1.0 | 1.4 |
| Eastern Africa | 1.4 | 1.7 | 1.8 | 1.9 | 1.1 | 1.6 |
| Southern Africa | 1.1 | 2.1 | 1.0 | 1.8 | 1.2 | 1.4 |
| N Africa / W Asia / Europe | 1.7 | 1.4 | 1.1 | 1.1 | 0.9 | 1.2 |
| Central Asia | | 1.0 | 1.6 | | 0.9 | 1.1 |
| South and SE Asia | 2.7 | 1.2 | 1.5 | 0.9 | 1.2 | 1.3 |
| Latin America and Carib. | 0.7 | 1.1 | 0.9 | 1.2 | 0.5 | 0.9 |
| All regions | 1.6 | 1.5 | 1.4 | 1.6 | 1.2 | 1.5 |

Displacement of births is much easier to identify than omission. Most of the displacement in Table 6.4 is shifting of birthdates backward across the boundary for the health questions, usually five years before the date of interview. Table 6.4 shows that the average level of displacement of births across all 192 surveys is 1.5 percent. Comparing DHS-2 with DHS-6, there is a reduction from 1.6 to 1.2 percent. The reduction was not continuous, but DHS-6 has the lowest level of all phases. Western Africa has the highest mean level of displacement, 2.1 percent overall and 2.0 percent in DHS-6. Latin America has the lowest mean level of displacement, 0.9 percent overall and 0.5 percent in DHS-6. All other regions are in a narrow range of 1.1 percent to 2.1 percent for all phases combined, and 0.9 percent to 2.0 percent in DHS-6.

Applying the approximation given earlier, on average, an estimate of the ratio of fertility 0-4 years before the survey to fertility 5-9 years before the survey has a potential downward bias of about 6 percent for all 192 surveys (4*1.5=6) and about 5 percent for DHS-6 surveys (4*1.2=4.8). This is just a rough guide; the bias will be less if the window for the health questions is something other than five years and if the interval for the fertility rates is different from the interval for the health questions. In general, the interval “0-4 years ago” will extend only part way into the fifth calendar year preceding fieldwork, reducing the bias. The bias would be greater for surveys with higher than average levels of displacement.

## 6.5. Estimated Probability of Omission of Deaths

**Table 6.5. Mean estimated probability of omission of an under-5 death, by region and phase of DHS, 192 DHS surveys and births in the 10 years before each survey**

| | Phase of DHS | | | | | |
|---|---|---|---|---|---|---|
| **Region** | **2** | **3** | **4** | **5** | **6** | **All phases** |
| Western Africa | 3.3 | 2.7 | 3.1 | 3.1 | 2.6 | 2.9 |
| Middle Africa | 2.9 | 2.3 | 1.4 | 5.1 | 6.7 | 3.8 |
| Eastern Africa | 5.1 | 6.0 | 4.2 | 5.5 | 3.2 | 4.7 |
| Southern Africa | 8.7 | 13.8 | 6.2 | 8.0 | 3.4 | 7.7 |
| N Africa / W Asia / Europe | 4.6 | 4.4 | 5.1 | 6.3 | 9.8 | 5.7 |
| Central Asia | | 9.6 | 1.7 | | 15.7 | 9.2 |
| South and SE Asia | 5.2 | 5.8 | 6.5 | 7.1 | 5.0 | 6.1 |
| Latin America and Carib. | 8.4 | 6.0 | 8.5 | 5.8 | 5.3 | 6.7 |
| All regions | 5.5 | 5.4 | 5.0 | 5.6 | 4.8 | 5.3 |

Omission of under-5 deaths in this report, as described in Section 5.9, includes omission of births that resulted in deaths. In effect, a death rate is a ratio of deaths to births, and omissions will remove some deaths in the numerator and the corresponding births in the denominator. The means of the estimated probabilities of omission, in each combination of region and phase, are given in Table 6.5.

The overall mean probability of omitting a death is 5.3 percent, nearly three times the level of omission of a birth, 1.8 percent. Table 6.3 showed that the estimated probability of omitting a birth declined from DHS-2 to DHS-6, from 2.0 percent to 1.5 percent, a decline of about 25 percent; by contrast, Table 6.5 suggests that the estimated probability of omitting a death showed a slight increase, from 5.5 percent in DHS-2 to 5.6 percent in DHS-5, but a net decline to 4.8 percent in DHS-6. That is, for both births and deaths there was an overall decline in omission from DHS-2 to DHS-6, but it was not steady. Most regions showed a drop in omission from DHS-5 to DHS-6; the exceptions are Middle Africa, North Africa / West Asia / Europe, and Central Asia. The level of 15.7 percent for Central Asia in DHS-6 is the highest figure in Table 6.5, much higher than the 1.7 percent for DHS-4 (there were no surveys in that region in DHS-5). We do not believe that the actual omission of deaths reached such high levels in those surveys, but further investigation will be required to determine why the symptoms of omission of deaths reached such high levels for the surveys in Central Asia.

The low estimates of omission of deaths in Western Africa, averaging only 2.6 percent in DHS-6, are unexpected. There will be further discussion of omission in Chapter 7.

### 6.6. Estimated Probability of Displacement of Deaths

**Table 6.6. Mean estimated probability of displacement of an under-5 death, by region and phase of DHS, 192 DHS surveys and births in the 10 years before each survey**

| | Phase of DHS | | | | | |
|---|---|---|---|---|---|---|
| **Region** | **2** | **3** | **4** | **5** | **6** | **All phases** |
| Western Africa | 6.4 | 4.9 | 4.3 | 7.0 | 5.9 | 5.6 |
| Middle Africa | 3.5 | 4.2 | 4.7 | 4.0 | 4.6 | 4.3 |
| Eastern Africa | 5.3 | 5.3 | 5.4 | 5.5 | 4.9 | 5.3 |
| Southern Africa | 6.3 | 4.8 | 3.1 | 6.8 | 3.7 | 5.0 |
| N Africa / W Asia / Europe | 8.5 | 7.2 | 5.3 | 4.2 | 5.5 | 5.8 |
| Central Asia | | 3.7 | 4.3 | | 2.2 | 3.5 |
| South and SE Asia | 5.7 | 4.9 | 3.4 | 3.7 | 3.3 | 4.1 |
| Latin America and Carib. | 3.8 | 3.8 | 3.1 | 3.4 | 2.5 | 3.4 |
| All regions | 5.7 | 4.8 | 4.4 | 4.8 | 4.4 | 4.7 |

Table 6.6 gives estimates of the probability that a death has been displaced, by region and phase of DHS. The denominator of each estimate includes the deaths that were omitted.

Displacement of deaths arises from the displacement of births across the boundary for the child health questions, plus the heaping of deaths at age 12 months and at age 7 days. Table 6.6 gives the means of the estimated probabilities of displacement of deaths for the surveys within each combination of region and phase. The mean probability is 4.7 percent, which is close to the estimated mean probability of omission of deaths (5.3 percent) and substantially larger than the mean probability of displacement of births (1.5 percent). DHS-6 had the lowest mean probability of displacement of deaths, 4.4 percent, and DHS-2 the highest mean, 5.7 percent; again the decline from DHS-2 to DHS-6 was not monotonic. The range across all phases of DHS is relatively small. Except for Latin America and the Caribbean, Central Asia is the region with the lowest level for displacement of deaths; in Table 6.5 it was the region with the highest level of omission. In DHS-5 and DHS-6, Western Africa is the region with the highest levels of displacement of deaths.

# 7. Alternative Approaches to Omission

The three types of omission, or potential omission, identified in this report were essentially based on comparing one count with another count and imposing a criterion for what the ratio of those counts should be. For the first two types, a comparison was made between the numbers of boys and girls that were enumerated, and the criterion was a hypothesized ratio of these numbers. For the third type, a comparison was made between the numbers of neonatal and infant deaths, and the criterion was a hypothesized proportion that depended on the level of infant mortality. For all three types of omission, the hypothesized values for the criteria were based on more broadly observed empirical regularities found with some consistency in other data. These include the sex ratio at birth, the sex ratio of neonatal deaths, and the empirical correspondence between the proportion of infant deaths that are neonatal and the infant mortality rate.

The strategy of comparing one group with another group is limited, at best, to identifying relative omission. It cannot identify more general types of omission that affect both the comparison group and the reference group. It is likely that the DHS birth histories are sometimes affected by more general omission. An example that is particularly challenging to identify is the possible omission of children who died before the survey, with only a weak relationship (if any) to the age of the child at death, the sex of the child, and when the death occurred.

The Benin 2011-12 DHS survey shows evidence of this type of omission. The main report on this survey (INSAE 2013, pp. 219-223) explicitly acknowledges the omission of children who died. The report compares the levels of infant and child mortality coming from this survey with the levels in the preceding survey, conducted in 2006. For the periods prior to the 2006 survey, the mortality estimates from the 2011-12 survey are about 40 percent below the estimates from the 2006 survey. In a similar comparison of fertility rates, the estimates are somewhat lower in the more recent survey, although by a much smaller percentage than the apparent deficit in the infant and child mortality rates.

The evidence suggests that in the 2011-12 Benin DHS survey there was substantial omission of children who did not survive to the date of interview. These omissions did not have a particular tendency to be children who died when very young, because the reported balance between neonatal and infant mortality is almost exactly what we would expect it to be, given the observed level of infant mortality. The omissions also do not appear to be strongly related to how long ago the child died, because the deficits for 5-9 and 10-14 years before the survey appear to be about the same.

By definition, across-the board omission of children who died falls outside the framework of ratios or proportions and criterion values described above. This framework can only deal with omission in one category relative to another category. Indeed, in comparing two successive surveys, such as those from Benin, the rough estimate made above that “the mortality estimates from the 2011-12 survey are about 40 percent below the estimates from the 2006 survey” can only identify relative omission. It is possible that the Benin 2006 DHS also had overall omission, and the evidence only goes so far as to suggest that the 2011-12 survey had *more* omission than the 2006 survey. The level of omission in either survey remains uncertain.

Two potential strategies to identify across-the-board omission of children who died will be briefly described. The first potential strategy would be based on the distribution of lengths of birth intervals. Suppose that children who survived to the date of interview are reported completely, but children who died tend to be omitted. The evidence in the birth histories will be that some birth intervals will be longer than they should be. These will appear as unexpectedly long intervals between two recorded births, within which a live birth (or more than one) occurred and the child subsequently died.

To simulate this pattern of omission, we analyzed in detail specific surveys that appear to have high quality reporting of dates in the birth histories, no evidence of omission, and only a moderate level of child mortality. In such a survey the omission of children who died is simulated by randomly masking such children and recalculating birth intervals between the remaining live births. The goal of the simulation exercise was to devise a procedure to recover the masked births, at least to the degree of specifying a probability that a long interval actually included an omitted birth.

An important complication, if such an "unmasking" procedure is applied to real birth histories as distinguished from artificially masked birth histories, is that survivorship is not independent of the lengths of the actual subsequent and preceding birth intervals. Children who died tend to have died when very young, reducing the duration of breastfeeding and postpartum amenorrhea and the interval to the next conception. Also, a short preceding birth interval will tend to increase the chance that the child will die (see, for example, Rutstein and Winter 2013). For both of these reasons, the actual omission of a child who died will extend the interval somewhat but typically will not double its length.

In the collection of birth histories, long intervals between births have long been interpreted as possible evidence of omission. DHS data collection includes probes for possible omission if the stated birth interval is unexpectedly long.

Various models were used in an attempt to estimate the probability that an interval includes an unreported child who died, based on the length of the interval between two recorded births. Unfortunately, these efforts involve a confounding of differences in levels of omission with genuine differences in child mortality, and are also sensitive to the choice of a reference survey or subsample. There is an unacceptable level of uncertainty and dependence on assumptions in the application of this approach, and no results from it are included in this report.

A second potential strategy to identify the broad omission of children who died is an extension of the Hill and Choi (2006) pattern that was the basis of our estimates of the omission of neonatal deaths. They found a similar relationship for early neonatal mortality. The two patterns found by Hill and Choi suggest a possible extrapolation, to an empirical relationship between the ratio of infant mortality to under-5 mortality and the log of under-5 mortality.

We investigated this model with data available from the United Nations Inter-agency Group for Child Mortality Estimation (UN-IGME; http://www.data.unicef.org/child-mortality/under-5). There is indeed some empirical support for it, but the relationship is not strong enough to justify applying it. Note also that, although such a model could potentially identify omission of infant deaths, it still does not address across-the-board omission that affects the full under-5 age range.

It may be possible to develop a more general model for omission based entirely on internal evidence, but we have found it to be difficult to generate defensible estimates of omission using just the birth histories in a single survey. The three specific types of relative omission discussed in this report are certainly incomplete and imperfect, sometimes producing estimates that are too high and sometimes too low, with no possibility of validation. For this reason we have often referred to "potential" omission.

# 8. Discussion and Conclusions

This report has been methodological in two senses. First, it is one of a long series of assessments of the quality of DHS data. It has focused on the birth histories, which are the basis of the DHS estimates of fertility and child mortality, which are in turn probably the most important of all the rates that DHS produces. The report is a companion to another DHS report by Schoumaker (2014) that focuses directly on the fertility rates and gets additional leverage from adjacent surveys in the same country. It is also related to a DHS report by Ahmed et al. (2014) that assesses the sibling histories and the maternal mortality estimates. This report was preceded by another one on date and age reporting by Pullum (2006) that included the birth histories but with much less detail. These kinds of assessments of large numbers of surveys at a time should not obscure the fact that every single DHS survey receives careful assessment during the phases of data collection, editing, imputation, the calculation of indicators and the preparation of the main report. The main report on every DHS survey includes an appendix on data quality.

There is a second sense in which this report has been methodological: it has attempted to advance the procedures for carrying out an assessment of birth history data. We have developed a framework to identify potential omission and displacement of dates that has included these key steps:

1. Specify criteria that should be satisfied by "true" or correct measurements.
2. Examine each survey's deviations from the criteria, as indicators of data quality.
3. Adjust the birth histories to match the criteria, using a reweighting procedure, one criterion at a time and then all criteria together, iterating until convergence.
4. Examine the differences between the original births and deaths, and the reweighted births and deaths, to obtain estimates of the numbers of omissions and displacements.

We believe that this framework is new. In addition, we examined the completeness of birth dates (b10) and the flagging of ages at death (b13), two longstanding indicators of the quality of the birth histories.

We emphasize that the choices of specific criteria—three for omission and three for displacement—are distinct from the framework. The entire exercise could be repeated with fewer criteria, more criteria, different criteria, or different reference values for the omission criteria. There are other ways in which some of the procedures could be modified while retaining the basic framework. For example, instead of births in the past 10 years, it might be possible to use a shorter or longer interval. It would be possible to model the probabilities of omission and displacement as functions of the time since the event. The procedure could be applied separately to urban and rural populations or could be expanded to include covariates such as age of the mother, her level of education, interviewer or supervisor codes, etc. The results could change substantially with such modifications or enhancements.

Because the basic framework could include so many possible variations, we do not recommend that the findings reported in Chapters 4 and 5, especially for omission, be viewed as measures of the data quality for specific surveys. Rather, large deviations from the criteria and a large calculated impact of deviations from the criteria should be interpreted as symptoms or warnings, rather than as a diagnosis of poor data quality in specific surveys. The means provided in Chapter 6, structured by geographical region and phase of DHS, are less affected by the shortcomings of the indicators.

In medical terminology, some of the surveys with relatively large deviations from the omission criteria in Chapters 4 and 5 are false positives. The deviations may suggest omission, but those deviations may be due to sampling error or to genuine country-specific differences between the hypothesized value of the criterion and the actual value for the country. For example, the model for the proportion of infant deaths that are neonatal may not apply to a specific country. If the model does not apply, then even a well-executed survey will show a deviation that is not due to poor data quality. Successive surveys in such a country will repeatedly give spurious evidence of poor quality.

By the same token, some of the surveys with negligible deviations from the omission criteria are false negatives. This happens when the country matches the criteria but the criteria are unable to capture the pattern of omission that actually characterizes the survey. For example, there may be omission of child deaths that is not captured by our limited set of indicators. As described in Chapter 7, the criteria cannot capture across-the-board omission of children who died.

We would have expected some specific DHS surveys to show clear evidence of omission, in particular, because of their high rates of incompleteness (b10) and flagging (b13), but the criteria did not consistently indicate omission in those surveys. These false negatives include some of the surveys conducted in West Africa. As a specific case, the Nigeria 1999 DHS is well known to have been a survey with poor quality. Other reports by Pullum (2006) and Schoumaker (2014) have identified serious misreporting in this survey. This was a survey with minimal involvement by DHS staff during data collection. The poor quality of the Nigeria 1999 survey is a demonstration of the added value of DHS technical assistance during data collection. We would have expected the survey to fare badly in terms of omission, but it did not. Other surveys, such as the Benin 2011-12 survey, singled out in Chapter 7, are also clearly false negatives for omission.

The indicators in this report can be structured into six categories: incompleteness of birthdates; flagging of age at death; omission of births; displacement of births; omission of deaths; and displacement of deaths. Each of these categories has one indicator, except that omission of deaths and displacement of deaths have two indicators each. It could plausibly be hypothesized that there is an underlying or latent structure to the misreporting of events and dates, which would lead the indicators of incompleteness, flagging, omission, and displacement to be strongly and positively correlated with one another. Whatever the sources of genuine data quality may be, ranging across characteristics of the survey design, the wording and flow of the questionnaire, the training and supervision of interviewers, or the respondents' cooperation and knowledge of dates and ages, we would expect the indicators of data quality to be correlated. It is therefore surprising that when we examined all possible correlations—although only some of them were discussed—we did not find evidence of underlying dimensions or factors. We only found one significantly positive correlation. The correlation between incompleteness of birthdate (b10) and potential displacement of births was 0.33 ($p<.0001$). The correlation between omission and displacement of births was significantly negative (-.15 with $p<.05$), and the correlation between omission and displacement of deaths was significantly negative (-.24 with $p<.0001$). We would have expected the symptoms of omission and displacement to be positively correlated, if both of them are indeed indicators of data quality. Incompleteness of birthdate and flagging of age at death are not significantly correlated across the surveys.

There are at least three possible interpretations of the lack of strong and positive correlations among the indicators. One possibility is that data quality is multi-faceted and complex, and cannot be distilled into underlying dimensions or factors. Another possibility is that the indicators used in this study are not doing what they were intended to do—that is, to measure data quality. A third possibility is that most DHS surveys are of high quality, and for such surveys deviations from the criteria occur essentially at random. Thus, even if there is an underlying or latent structure to data quality, and even if we have specified indicators that reflect that structure, it could happen that the random component of the deviations from the indicators has swamped the systematic component and has severely attenuated an expected strong and positive association among the indicators. We suggest that all three interpretations are plausible, or at least cannot be rejected.

In terms of the DHS surveys as a whole, the data quality of the birth histories appears to be very high. Incompleteness of birthdates has declined over the phases of DHS and reached its lowest level in DHS-6.

In most surveys there is little evidence of omission of the births of surviving children. Displacement of birthdates, outside the window for the health questions, has been serious but has declined in importance in most regions, reaching its lowest point during DHS-6. Displacement can have an important impact on estimates of fertility change if, say, the boundary between the most-recent time interval and the next-most-recent time interval coincides exactly with the boundary for the health questions. If those two intervals are 0-4 and 5-9 years before the survey, respectively, and the boundary for the health questions is January of the fifth calendar year preceding the first year of fieldwork, then the boundary for the intervals and the boundary for the health questions will be only a few months apart, and many surveys will show some spurious evidence of a decline in fertility, or exaggerated evidence of a decline.[22]

The data on child deaths, in general, are not as good as the data on births. Heaping on age 12 months at death is found in many surveys. This type of displacement will cause a downward bias in infant mortality (age 0 years) and an upward bias in child mortality (age 1-4 years), but will have virtually no impact on the under-5 mortality rate. Displacement across the boundary for the health questions will have only a minor effect on the estimation of trends in child survival, because the births and deaths are displaced together and ratios of deaths to births are largely unaffected.

Omission of child deaths is believed to be the greatest potential weakness in the birth histories, although it is probably serious only for a small fraction of DHS surveys. It is probably also the main cause of missing births, although the effect on fertility rates is much smaller than the effect on mortality rates. Such omission is difficult to identify on the basis of internal evidence within a single survey. Our indicators of such omission are inadequate, but they suggest that it has been an issue during all phases of DHS and, in some regions, may even have been higher in DHS-6 than in earlier phases. The findings are consistent with concerns about under-5 mortality estimates from some recent surveys in sub-Saharan Africa. Nevertheless, the indicators of omission used in this report have not successfully identified the specific surveys that have the greatest *prima facie* evidence of omission coming from a comparison of under-5 mortality rates in successive surveys in recent DHS main reports or a comparison of DHS estimates with other estimates. A comprehensive search for omission seems to require a comparison of multiple surveys or other sources—although even in that context, the identification of *relative* omission may be as much as is possible.

It is recommended that further analysis be carried out on some of the specific surveys that appear to have the highest values of the indicators. In this report we have tended to discount some of the high values, as being attributable to inappropriate standards for the criteria or to small sample sizes, but some further investigation should be undertaken.

We recommend further analysis of repeated appearances of high values of an indicator in two successive surveys of the same country. It is not clear whether the repeated appearances are due to an inappropriate reference value for a criterion or to repeated problems with implementation of the surveys.

Another recommendation is that the basic framework or methodology be reapplied with alternative criteria, alternative reference values, and covariates such as interviewer characteristics. If the framework itself can be validated, at least in a probabilistic sense, then it could become a tool for the assessment of data quality in specific surveys after they are completed, or even a tool to provide feedback as the birth histories are being collected.

[22] The boundary for the health questions is a fixed date, as stated, but the boundary between 0-4 and 5-9 years before the survey will fluctuate according to the mother's date of interview. The latter boundary will usually be several calendar months after the health boundary, a factor that will tend to reduce the effect of birth displacement on the fertility rates.

# References


Alkema, L., F. Chao, D. You, J. Pederson, and C.C. Sawyer. 2014. “National, Regional, and Global Sex Ratios of Infant, Child, and Under-5 Mortality and Identification of Countries with Outlying Ratios: A Systematic Assessment.” *Lancet Global Health* 2: 521-30.

Arnold. F. 1990. “Assessment of the Quality of Birth History Data in the Demographic and Health Surveys.” In *An Assessment of DHS-I Data Quality*, DHS Methodological Reports No. 1, page 83-111. Columbia, Maryland, USA: Institute for Resource Development and Macro Systems Inc.

Ayeni, O. 1975. “Sex Ratio of Live Births in South-Western Nigeria.” *Annals of Human Biology* 2: 137-41.

Azeez, M.A., A. Akinboro, and A.A. Bakare. 2007. “Human Sex Ratio at Birth in Southwest Nigeria.” *Indian Journal of Human Genetics* 13(2): 59-64.

Basu, A. 2005. “Extended Generalized Linear Models: Simultaneous Estimation of Flexible Link and Variance Functions.” *The Stata Journal* 5: 501-16.

Chahnazarian, A. 1988. “Determinants of the Sex Ratio at Birth: Review of Recent Literature.” *Social Biology* 35(3/4): 214-35.

Chidambaram, V.C., and T.W. Pullum. 1981. “Estimating Fertility Trends from Retrospective Birth Histories: Sensitivity to Imputation of Missing Dates.” *Population Studies* 35: 307-20.

Croft, T. 1991. “DHS Editing and Imputation.” Available online as “DHSG3 under DHS Questionnaires and Manuals,” at www.dhsprogram.com

Deming, W.E. 1943. *The Statistical Adjustment of Data*. Mineola, New York, USA: Dover Publications.

Erickson, D. 1976. “The Secondary Sex Ratio in the United States 1969-71: Association with Race, Parental Ages, Birth Order, Paternal Education and Legitimacy.” *Annals of Human Genetics* 40: 205-12.

Garenne, M. 2004. “Sex Ratios at Birth in Populations of Eastern and Southern Africa.” *South African Journal of Demography* 9(1): 91-6.

Garenne, M. 2008. “Variations in the Sex Ratio at Birth in African Demographic Surveys.” *Human Biology* 80(5): 473-82.

Guillot, M., P. Gerland, F. Pelletier, and A. Saabneh. 2012. “Child Mortality Estimation: A Global Overview of Infant and Child Mortality Age Patterns in Light of New Empirical Data.” *PLOS Medicine* 9(8): e1001299.

Helleringer, S., G. Pison, A.M. Kanté, G. Duthé, and A. Andro. 2014. “Reporting Errors in Siblings’ Survival Histories and Their Impact on Adult Mortality Estimates: Results from a Record Linkage Study in Senegal.” *Demography* 51(2): 387-411.

Hill, K., and Y. Choi. 2006. “Neonatal Mortality in the Developing World.” *Demographic Research* 14: 429-52.

Institut National de la Statistique et de l’Analyse Économique (INSAE), et ICF International. 2013. *Enquête Démographique et de Santé du Bénin 2011-2012*. Calverton, Maryland, USA: INSAE, et ICF International.

James, W.H. 1984. “The Sex Ratios of Black Births.” *Annals of Human Biology* 11(1): 39-44.

Little, R.J.A. 1993. “Post-Stratification: A Modeler’s Perspective.” *Journal of the American Statistical Association* 88: 1001-12.

Marckwardt, A.M., and S.O. Rutstein, 1996. *Accuracy of DHS-II Demographic Data: Gains and Losses in Comparison with Earlier Surveys.* DHS Working Papers No. 19. Calverton, Maryland, USA: Macro International.

Navara, K.J. 2009. “Humans at Tropical Latitudes Produce More Females.” *Biology Letters* 5: 524-27.

Pullum, T.W. 2006. *An Assessment of Age and Date Reporting in the Demographic and Health Surveys, 1985-2003.* DHS Methodological Reports No. 5. Calverton, Maryland, USA: Macro International.

Pullum, T.W., and S.L. Stokes. 1997. “Identifying and Adjusting for Recall Error, with Application to Fertility Surveys.” In *Survey Measurement and Process Quality*, edited by L.E. Lyberg, P.P. Biemer, M. Collins, E.D. de Leeuw, C. Dippo, N. Schwarz, and D. Trewin, page 711-32. Hoboken, New Jersey, USA: John Wiley and Sons.

Ruder, A. “Paternal Factors Affect the Human Secondary Sex Ratio in Interracial Births.” *Human Biology* 58(3): 357-66.

Schoumaker, B. 2014. *The Quality and Consistency of Fertility Estimates from DHS Birth Histories.* DHS Methodological Reports No. 12. Rockville, Maryland, USA: ICF International.

Sullivan, J.M., G.T. Bicego, and S.O. Rutstein. 1990. “Assessment of the Quality of Data Used for the Direct Estimation of Infant and Child Mortality in the Demographic and Health Surveys.” In *An Assessment of DHS-I Data Quality*, DHS Methodological Reports No. 1, page 115-38. Columbia, Maryland, USA: Institute for Resource Development and Macro Systems Inc.

Tarver, J.D., and C. Lee. 1968. “Sex Ratio of Registered Live Births in the United States 1942-1963.” *Demography* 5(1): 374-81.

Teitelbaum, M.S. 1971. “Male and Female Components of Perinatal Mortality: International Trends 1901-1963.” *Demography* 8(4): 541-48.

Ulizzi, L., and L.A. Zonta, 2002. “Sex Differential Patterns in Perinatal Deaths in Italy.” *Human Biology* 74(6): 879-88.

Visaria, P. 1967. “Sex Ratio at Birth in Territories with Relatively Complete Registration.” *Eugenics Quarterly* 14: 132-42.

# Appendix Indicators of Omission or Displacement for All Surveys

**Table A1. Percentage of births with incomplete birthdate, percentage of deaths with flagged age at death, deviations from the criteria for omission, and relative deviations from the criteria for displacement, all 192 surveys and births in the past 10 years**

| Survey | Region | Phase of DHS | Percentages of: Births with incomplete birthdate | Percentages of: Deaths with flagged age at death | Deviation for sex ratio at birth | Deviation for sex ratio of neonatal deaths | Deviation of percent of infant deaths that are neonatal | Relative deviation (x100) for deaths at 12 months | Relative deviation (x100) for deaths at 7 days | Relative deviation (x100) for births in year before boundary |
|---|---|---|---|---|---|---|---|---|---|---|
| Albania 2008-09 | N Africa / W Asia / Europe | 5 | 0 | 2 | 5.8 | | -6.4 | -38.3 | 79.9 | 4.1 |
| Armenia 2000 | N Africa / W Asia / Europe | 4 | 0 | 3 | 10.0 | -2.7 | 7.1 | -59.3 | 50.5 | 8.3 |
| Armenia 2005 | N Africa / W Asia / Europe | 5 | 0 | 32 | 17.0 | | -16.0 | 251.3 | 69.3 | 8.8 |
| Armenia 2010 | N Africa / W Asia / Europe | 6 | 0 | 0 | 8.7 | | 39.5 | 161.4 | 335.3 | 6.8 |
| Azerbaijan 2006 | N Africa / W Asia / Europe | 5 | 0 | 7 | 14.3 | -2.9 | -88.6 | -33.7 | 144.2 | 8.9 |
| Bangladesh 1993-94 | South and SE Asia | 3 | 0 | 1 | -1.0 | -1.0 | 18.4 | 173.1 | -5.6 | 2.0 |
| Bangladesh 1996-97 | South and SE Asia | 3 | 0 | 5 | -5.7 | -2.2 | 28.2 | 84.0 | 15.9 | 8.1 |
| Bangladesh 1999-00 | South and SE Asia | 4 | 1 | 2 | -0.7 | -1.5 | 23.8 | 160.8 | 22.9 | 6.3 |
| Bangladesh 2004 | South and SE Asia | 4 | 0 | 2 | -1.1 | -0.9 | 15.6 | 77.1 | 32.8 | 3.5 |
| Bangladesh 2007 | South and SE Asia | 5 | 0 | 3 | -2.9 | -2.6 | 32.4 | 63.0 | 89.6 | 3.4 |
| Bangladesh 2011 | South and SE Asia | 6 | 0 | 0 | 0.7 | -0.9 | 9.7 | 32.4 | 19.9 | 6.0 |
| Benin 1996 | Western Africa | 3 | 36 | 3 | -0.8 | -2.9 | 24.3 | 9.7 | 27.5 | 5.3 |
| Benin 2001 | Western Africa | 4 | 33 | 4 | -1.0 | -2.8 | 35.7 | 99.3 | 77.2 | 8.5 |
| Benin 2006 | Western Africa | 5 | 10 | 7 | -0.5 | -1.9 | 24.1 | 101.7 | 37.1 | 18.8 |
| Benin 2011-12 | Western Africa | 6 | 26 | 6 | 4.4 | -0.6 | 13.9 | 164.7 | 8.5 | 13.1 |
| Bolivia 1994 | Latin America and Carib. | 3 | 2 | 1 | -0.6 | -4.4 | 41.4 | 129.8 | -22.0 | 8.6 |
| Bolivia 1998 | Latin America and Carib. | 3 | 2 | 2 | -0.3 | -0.9 | 17.1 | 82.9 | -43.4 | 6.2 |
| Bolivia 2003 | Latin America and Carib. | 4 | 1 | 1 | -1.7 | -1.4 | 27.0 | 65.5 | -3.4 | 4.1 |
| Bolivia 2008 | Latin America and Carib. | 5 | 0 | 2 | -0.6 | -1.3 | 16.8 | 64.0 | 12.2 | 6.1 |
| Brazil 1991 | Latin America and Carib. | 2 | 1 | 24 | 5.2 | -4.5 | -19.4 | 98.2 | -14.4 | -1.2 |
| Brazil 1996 | Latin America and Carib. | 3 | 2 | 8 | 0.0 | -5.9 | 48.6 | 9.3 | -17.9 | 6.3 |
| Burkina Faso 1993 | Western Africa | 2 | 22 | 1 | 1.0 | 1.0 | 19.0 | 222.4 | 1.6 | 21.2 |
| Burkina Faso 1998-99 | Western Africa | 3 | 36 | 6 | 0.5 | 0.8 | -2.3 | 91.7 | 11.6 | 12.3 |
| Burkina Faso 2003 | Western Africa | 4 | 4 | 6 | 0.1 | -0.1 | 19.6 | 164.6 | 25.9 | 13.1 |
| Burkina Faso 2010 | Western Africa | 6 | 1 | 2 | 1.4 | -1.9 | 24.4 | 50.2 | 21.7 | 14.5 |

**(Continued...)**

**Table A1. – *Continued***

| Survey | Region | Phase of DHS | Percentages of: Births with incomplete birthdate | Percentages of: Deaths with flagged age at death | Deviation for sex ratio at birth | Deviation for sex ratio of neonatal deaths | Deviation of percent of infant deaths that are neonatal | Relative deviation (x100) for deaths at 12 months | Relative deviation (x100) for deaths at 7 days | Relative deviation (x100) for births in year before boundary |
|---|---|---|---|---|---|---|---|---|---|---|
| Burundi 2010 | Eastern Africa | 6 | 1 | 0 | -1.2 | -0.5 | 17.5 | 66.6 | -17.2 | 3.4 |
| Cambodia 2000 | South and SE Asia | 4 | 0 | 7 | -1.2 | -2.4 | 16.8 | 73.7 | -23.8 | 21.7 |
| Cambodia 2005 | South and SE Asia | 5 | 1 | 5 | -3.1 | 0.2 | 6.7 | 167.6 | -5.3 | 0.1 |
| Cambodia 2010 | South and SE Asia | 6 | 1 | 2 | 1.2 | -0.9 | 8.7 | 184.2 | 40.4 | 6.2 |
| Cameroon 1991 | Middle Africa | 2 | 22 | 1 | -2.1 | -0.1 | 16.6 | 68.3 | 10.9 | 8.2 |
| Cameroon 1998 | Middle Africa | 3 | 5 | 8 | -5.2 | -1.7 | 8.5 | -9.2 | 26.6 | 9.9 |
| Cameroon 2004 | Middle Africa | 4 | 6 | 7 | -1.8 | 0.1 | 10.8 | 53.5 | 39.4 | 4.6 |
| Cameroon 2011 | Middle Africa | 6 | 3 | 7 | -2.4 | -1.8 | 27.2 | 58.5 | 47.4 | 10.9 |
| Central African Republic 1994-95 | Middle Africa | 3 | 6 | 1 | 1.3 | 0.1 | 9.9 | 53.1 | -30.5 | 6.1 |
| Chad 1996-97 | Middle Africa | 3 | 7 | 5 | -3.1 | 0.3 | 7.5 | 98.1 | -15.2 | 16.8 |
| Chad 2004 | Middle Africa | 4 | 1 | 1 | -0.1 | 1.1 | -1.5 | 104.2 | 36.7 | 17.3 |
| Colombia 1990 | Latin America and Carib. | 2 | 0 | 28 | -8.8 | | 47.1 | 192.8 | -24.2 | 1.4 |
| Colombia 1995 | Latin America and Carib. | 3 | 0 | 4 | 0.7 | -1.7 | 11.5 | 1.0 | 83.8 | 4.6 |
| Colombia 2000 | Latin America and Carib. | 4 | 0 | 0 | -0.3 | | -5.3 | -42.2 | -24.4 | 2.0 |
| Colombia 2005 | Latin America and Carib. | 5 | 0 | 0 | -1.7 | | 2.5 | 81.9 | 31.1 | 3.3 |
| Colombia 2010 | Latin America and Carib. | 6 | 0 | 11 | 2.1 | | -5.7 | -70.2 | 44.1 | 0.9 |
| Comoros 1996 | Eastern Africa | 3 | 26 | 1 | -2.4 | -1.0 | 20.9 | 194.9 | -16.7 | 11.5 |
| Congo (Brazzaville) 2005 | Middle Africa | 5 | 2 | 5 | -0.4 | -2.3 | 25.8 | 79.1 | -23.7 | 3.3 |
| Congo (Brazzaville) 2011-12 | Middle Africa | 6 | 1 | 1 | -1.8 | -1.4 | 19.8 | 118.7 | 5.5 | -1.0 |
| Congo Democratic Republic 2007 | Middle Africa | 5 | 2 | 9 | -3.3 | -1.2 | 10.3 | 0.7 | -25.9 | 4.1 |
| Cote d'Ivoire 1994 | Western Africa | 3 | 11 | 0 | -0.5 | 0.1 | 18.6 | 62.9 | -17.3 | 14.0 |
| Cote d'Ivoire 1998-99 | Western Africa | 3 | 11 | 4 | -3.0 | -1.1 | 9.7 | 97.1 | -4.3 | 15.3 |
| Cote d'Ivoire 2011-12 | Western Africa | 6 | 3 | 4 | 1.2 | -1.4 | -33.5 | 131.3 | 39.2 | 12.4 |
| Dominican Republic 1991 | Latin America and Carib. | 2 | 1 | 0 | 1.0 | -1.5 | -43.6 | 118.2 | 93.0 | 7.6 |
| Dominican Republic 1996 | Latin America and Carib. | 3 | 1 | 6 | 1.7 | -2.1 | 13.4 | -57.1 | 56.7 | 0.3 |
| Dominican Republic 1999 | Latin America and Carib. | 4 | 1 | 10 | -13.4 | -9.0 | 58.7 | -100.0 | -4.9 | -4.3 |
| Dominican Republic 2002 | Latin America and Carib. | 4 | 1 | 6 | 0.9 | -0.3 | 5.6 | 84.2 | 29.9 | -0.2 |
| Dominican Republic 2007 | Latin America and Carib. | 5 | 1 | 4 | 2.1 | -2.3 | 21.1 | -40.0 | 48.3 | 6.2 |
| Egypt 1992 | N Africa / W Asia / Europe | 2 | 7 | 0 | -1.5 | 0.5 | 11.1 | 236.7 | -57.3 | 3.2 |
| Egypt 1995 | N Africa / W Asia / Europe | 3 | 13 | 0 | 2.3 | -0.3 | 4.8 | 233.9 | -38.3 | 3.7 |

**(Continued...)**

**Table A1. – *Continued***

| Survey | Region | Phase of DHS | Percentages of: Births with incomplete birthdate | Percentages of: Deaths with flagged age at death | Deviation for sex ratio at birth | Deviation for sex ratio of neonatal deaths | Deviation of percent of infant deaths that are neonatal | Relative deviation (x100) for deaths at 12 months | Relative deviation (x100) for deaths at 7 days | Relative deviation (x100) for births in year before boundary |
|---|---|---|---|---|---|---|---|---|---|---|
| Egypt 2000 | N Africa / W Asia / Europe | 4 | 5 | 3 | -1.6 | 0.6 | 17.5 | 258.4 | -50.6 | 11.5 |
| Egypt 2005 | N Africa / W Asia / Europe | 5 | 4 | 2 | -0.4 | -0.3 | 3.8 | 81.6 | 10.9 | 8.1 |
| Egypt 2008 | N Africa / W Asia / Europe | 5 | 1 | 1 | 0.8 | | -56.1 | 158.4 | 3.0 | 16.7 |
| Ethiopia 2000 | Eastern Africa | 4 | 7 | 5 | 3.1 | -0.3 | 3.9 | 152.1 | 1.8 | 6.9 |
| Ethiopia 2005 | Eastern Africa | 5 | 1 | 9 | 1.7 | 0.8 | -0.1 | 214.8 | -5.2 | 16.3 |
| Ethiopia 2011 | Eastern Africa | 6 | 2 | 3 | 2.7 | 0.8 | -9.7 | 155.5 | 68.8 | 7.8 |
| Gabon 2000 | Middle Africa | 4 | 2 | 5 | -4.8 | 0.2 | -3.6 | 69.5 | 64.4 | 7.8 |
| Gabon 2012 | Middle Africa | 6 | 2 | 1 | -5.5 | -2.6 | 39.2 | 146.0 | -57.1 | 6.4 |
| Ghana 1993 | Western Africa | 3 | 15 | 1 | 0.2 | 0.1 | 19.7 | 233.3 | -1.4 | 22.9 |
| Ghana 1998 | Western Africa | 4 | 12 | 6 | -0.1 | -0.3 | 16.2 | 125.4 | -42.4 | 7.3 |
| Ghana 2003 | Western Africa | 4 | 2 | 12 | 2.6 | -1.6 | 17.3 | 192.5 | 1.4 | 13.6 |
| Ghana 2008 | Western Africa | 5 | 2 | 12 | 2.8 | -0.7 | 17.1 | 172.5 | 9.1 | 15.0 |
| Guatemala 1995 | Latin America and Carib. | 3 | 1 | 1 | -1.0 | -0.6 | 11.9 | 134.5 | 83.7 | 6.7 |
| Guinea 1999 | Western Africa | 4 | 53 | 11 | 3.0 | 0.0 | 6.5 | 247.6 | -24.9 | 6.4 |
| Guinea 2005 | Western Africa | 5 | 36 | 4 | 3.5 | 1.7 | 1.6 | 118.0 | 17.4 | 15.3 |
| Guinea 2012 | Western Africa | 6 | 0 | 0 | 1.0 | 1.4 | 9.4 | 213.5 | 28.1 | 15.9 |
| Guyana 2009 | Latin America and Carib. | 5 | 1 | 10 | -2.4 | -3.8 | 32.9 | -10.8 | 0.8 | 2.6 |
| Haiti 1994-95 | Latin America and Carib. | 3 | 2 | 4 | -2.8 | -1.2 | 15.3 | 9.8 | -48.7 | 11.7 |
| Haiti 2000 | Latin America and Carib. | 4 | 1 | 1 | -8.6 | -2.3 | 37.1 | 93.6 | -34.7 | 8.9 |
| Haiti 2005-06 | Latin America and Carib. | 5 | 0 | 2 | -6.7 | -2.9 | 38.6 | 52.8 | 7.6 | 8.6 |
| Haiti 2012 | Latin America and Carib. | 6 | 0 | 0 | -0.5 | -2.1 | 27.6 | 133.8 | 21.1 | 2.9 |
| Honduras 2005-06 | Latin America and Carib. | 5 | 0 | 2 | 1.1 | | 20.0 | 6.8 | 8.3 | 3.1 |
| Honduras 2011-12 | Latin America and Carib. | 6 | 0 | 0 | 3.5 | | 45.5 | -0.2 | -5.8 | 4.5 |
| India 1992-93 | South and SE Asia | 2 | 2 | 1 | 1.0 | -1.9 | 24.4 | 111.4 | 1.7 | 6.6 |
| India 1998-99 | South and SE Asia | 4 | 2 | 3 | 2.2 | -1.8 | 28.0 | 181.3 | 26.4 | 7.4 |
| India 2005-06 | South and SE Asia | 5 | 1 | 5 | 2.4 | -2.4 | 28.3 | 90.1 | 41.3 | 2.7 |
| Indonesia 1991 | South and SE Asia | 2 | 7 | 0 | 3.9 | -3.6 | 27.0 | 142.7 | -26.7 | 9.9 |
| Indonesia 1994 | South and SE Asia | 3 | 7 | 1 | 2.0 | 0.5 | -6.5 | 209.5 | -18.7 | 6.3 |
| Indonesia 1997 | South and SE Asia | 3 | 6 | 1 | 2.6 | -1.9 | 21.2 | 186.9 | -6.4 | 12.8 |
| Indonesia 2002-03 | South and SE Asia | 4 | 5 | 7 | 0.7 | -4.2 | 29.4 | -74.2 | 17.7 | 16.1 |

**(Continued...)**

**Table A1. – *Continued***

| Survey | Region | Phase of DHS | Percentages of: Births with incomplete birthdate | Percentages of: Deaths with flagged age at death | Deviation for sex ratio at birth | Deviation for sex ratio of neonatal deaths | Deviation of percent of infant deaths that are neonatal | Relative deviation (x100) for deaths at 12 months | Relative deviation (x100) for deaths at 7 days | Relative deviation (x100) for births in year before boundary |
|---|---|---|---|---|---|---|---|---|---|---|
| Indonesia 2007 | South and SE Asia | 5 | 4 | 7 | 3.7 | -1.5 | 12.8 | 212.6 | 6.4 | 7.2 |
| Indonesia 2012 | South and SE Asia | 6 | 2 | 9 | -1.3 | -1.9 | -10.4 | 16.0 | 29.4 | 4.4 |
| Jordan 1990 | N Africa / W Asia / Europe | 2 | 1 | 0 | -0.4 | -3.5 | 38.9 | 282.8 | 92.9 | 1.6 |
| Jordan 1997 | N Africa / W Asia / Europe | 3 | 0 | 2 | 1.6 | | -6.4 | 247.1 | 2.0 | -0.3 |
| Jordan 2002 | N Africa / W Asia / Europe | 4 | 1 | 2 | -2.2 | | 35.3 | 231.6 | 46.8 | 8.4 |
| Jordan 2007 | N Africa / W Asia / Europe | 5 | 0 | 2 | 0.4 | | 25.2 | 98.2 | -21.2 | 1.4 |
| Jordan 2012 | N Africa / W Asia / Europe | 6 | 0 | 0 | 1.3 | | 21.8 | 219.0 | 30.8 | -6.1 |
| Kazakhstan 1995 | Central Asia | 3 | 0 | 4 | -2.3 | -7.5 | -36.4 | -14.7 | -19.8 | 2.4 |
| Kazakhstan 1999 | Central Asia | 4 | 0 | 4 | 1.5 | 0.1 | 5.9 | 87.7 | 104.0 | 4.5 |
| Kenya 1993 | Eastern Africa | 3 | 6 | 1 | -5.7 | -3.2 | 39.6 | 195.9 | -59.3 | 9.7 |
| Kenya 1998 | Eastern Africa | 3 | 1 | 4 | 0.9 | -5.3 | 23.3 | 129.7 | 15.2 | 4.0 |
| Kenya 2003 | Eastern Africa | 4 | 4 | 4 | -0.3 | -1.1 | 19.8 | 126.9 | 27.2 | 2.5 |
| Kenya 2008-09 | Eastern Africa | 5 | 2 | 10 | 2.4 | -0.4 | 5.1 | 148.4 | -25.5 | 12.6 |
| Kyrgyz Republic 1997 | Central Asia | 3 | 0 | 2 | -0.9 | -2.5 | 20.0 | 75.9 | 45.3 | 3.9 |
| Kyrgyz Republic 2012 | Central Asia | 6 | 0 | 7 | 3.5 | | 39.0 | -45.6 | 35.0 | 0.7 |
| Lesotho 2004 | Southern Africa | 4 | 1 | 5 | 0.3 | -3.1 | 28.9 | 1.6 | 73.3 | 6.2 |
| Lesotho 2009 | Southern Africa | 6 | 0 | 12 | 1.6 | -1.1 | -28.8 | 64.0 | -33.5 | 5.0 |
| Liberia 2007 | Western Africa | 5 | 2 | 15 | 1.3 | -1.6 | 14.7 | 5.4 | 9.9 | 24.5 |
| Madagascar 1992 | Eastern Africa | 2 | 14 | 1 | 2.4 | -1.1 | 26.3 | 149.5 | -64.4 | 7.8 |
| Madagascar 1997 | Eastern Africa | 3 | 17 | 1 | 0.7 | 0.3 | 4.8 | 64.7 | -41.0 | 9.7 |
| Madagascar 2003-04 | Eastern Africa | 4 | 6 | 13 | -1.7 | -0.8 | -7.6 | 20.2 | -39.2 | 22.5 |
| Madagascar 2008-09 | Eastern Africa | 5 | 1 | 7 | 1.1 | -4.2 | 17.1 | 2.0 | -33.4 | 15.6 |
| Malawi 1992 | Eastern Africa | 2 | 3 | 1 | -0.1 | -3.8 | 43.6 | 90.8 | -45.3 | 9.5 |
| Malawi 2000 | Eastern Africa | 4 | 1 | 5 | -3.1 | -1.5 | 27.4 | 137.5 | -0.7 | 22.4 |
| Malawi 2004 | Eastern Africa | 4 | 1 | 12 | -2.2 | -1.6 | 18.9 | 89.7 | -6.8 | 15.3 |
| Malawi 2010 | Eastern Africa | 6 | 0 | 13 | -3.7 | 0.5 | 7.1 | 80.2 | -44.4 | 6.4 |
| Maldives 2009 | South and SE Asia | 5 | 3 | 3 | -1.5 | | 24.0 | 3.9 | 40.9 | 1.7 |
| Mali 1995-96 | Western Africa | 3 | 3 | 2 | -2.5 | -0.9 | 19.8 | 72.7 | 22.0 | 10.3 |
| Mali 2001 | Western Africa | 4 | 4 | 7 | 0.4 | 2.4 | 3.1 | 204.3 | -22.9 | 9.1 |
| Mali 2006 | Western Africa | 5 | 5 | 5 | 1.6 | 1.3 | 3.3 | 131.0 | 49.4 | 24.2 |

(Continued...)

**Table A1. – *Continued***

| Survey | Region | Phase of DHS | Percentages of: Births with incomplete birthdate | Percentages of: Deaths with flagged age at death | Deviation for sex ratio at birth | Deviation for sex ratio of neonatal deaths | Deviation of percent of infant deaths that are neonatal | Relative deviation (x100) for deaths at 12 months | Relative deviation (x100) for deaths at 7 days | Relative deviation (x100) for births in year before boundary |
|---|---|---|---|---|---|---|---|---|---|---|
| Moldova 2005 | N Africa / W Asia / Europe | 5 | 0 | 27 | 1.0 | | -9.3 | | 45.0 | 1.4 |
| Morocco 1992 | N Africa / W Asia / Europe | 2 | 1 | 1 | 0.1 | -0.9 | 12.0 | 44.7 | -15.8 | -0.3 |
| Morocco 2003-04 | N Africa / W Asia / Europe | 4 | 5 | 6 | -3.6 | 0.0 | 6.3 | 81.1 | -24.2 | 4.1 |
| Mozambique 1997 | Eastern Africa | 3 | 27 | 1 | -2.3 | -2.9 | 39.5 | 174.5 | -34.3 | 20.8 |
| Mozambique 2003 | Eastern Africa | 4 | 2 | 4 | -5.8 | -3.9 | 44.8 | 197.4 | -17.3 | 13.5 |
| Mozambique 2011 | Eastern Africa | 6 | 2 | 2 | -2.0 | 0.0 | 12.0 | 200.4 | 18.4 | 7.1 |
| Namibia 1992 | Southern Africa | 2 | 5 | 3 | -6.1 | -1.9 | 28.6 | 154.0 | 20.0 | 4.6 |
| Namibia 2000 | Southern Africa | 4 | 2 | 9 | -3.7 | -2.2 | -27.0 | 68.5 | -42.7 | 9.4 |
| Namibia 2006-07 | Southern Africa | 5 | 1 | 9 | -0.4 | 0.5 | -26.0 | 214.3 | 21.5 | 11.5 |
| Nepal 1996 | South and SE Asia | 3 | 0 | 1 | -0.8 | -1.1 | 13.0 | 5.8 | 12.9 | 0.8 |
| Nepal 2001 | South and SE Asia | 4 | 0 | 1 | -4.8 | -2.0 | 28.9 | 37.9 | 30.2 | 0.1 |
| Nepal 2006 | South and SE Asia | 5 | 0 | 6 | 3.5 | -4.0 | 34.9 | -11.6 | 35.7 | 4.1 |
| Nepal 2011 | South and SE Asia | 6 | 0 | 1 | 2.5 | -3.1 | 30.9 | 15.0 | 143.3 | 4.3 |
| Nicaragua 1998 | Latin America and Carib. | 3 | 1 | 3 | -5.5 | -6.9 | 29.6 | 73.0 | 23.8 | -6.5 |
| Nicaragua 2001 | Latin America and Carib. | 4 | 1 | 2 | 1.0 | -8.7 | 28.0 | 90.7 | -0.7 | 11.5 |
| Niger 1992 | Western Africa | 2 | 31 | 7 | 2.0 | -3.7 | 41.0 | 81.6 | -30.4 | 11.2 |
| Niger 1998 | Western Africa | 3 | 3 | 0 | 0.6 | -0.4 | 6.3 | 45.7 | -7.2 | 9.0 |
| Niger 2006 | Western Africa | 5 | 11 | 3 | 0.4 | 1.9 | 26.4 | 276.8 | 19.3 | 25.6 |
| Niger 2012 | Western Africa | 6 | 11 | 3 | 0.1 | 3.7 | -7.7 | 276.5 | 13.7 | 27.4 |
| Nigeria 1990 | Western Africa | 2 | 12 | 4 | -0.4 | -0.3 | 26.9 | 147.6 | 16.1 | 18.2 |
| Nigeria 1999 | Western Africa | 4 | 15 | 26 | 1.5 | -2.3 | 22.3 | 67.4 | 11.0 | 12.0 |
| Nigeria 2003 | Western Africa | 4 | 7 | 11 | 1.3 | -1.5 | 14.6 | 39.1 | 15.2 | -0.4 |
| Nigeria 2008 | Western Africa | 5 | 2 | 17 | 0.9 | -1.6 | 19.8 | 91.3 | 31.1 | 11.0 |
| Pakistan 1990-91 | South and SE Asia | 2 | 8 | 6 | -0.6 | -0.5 | 14.4 | 67.3 | 19.7 | 31.8 |
| Pakistan 2006-07 | South and SE Asia | 5 | 17 | 4 | 4.5 | -1.1 | 19.2 | 227.8 | 27.1 | 13.0 |
| Paraguay 1990 | Latin America and Carib. | 2 | 0 | 1 | 1.4 | -3.0 | 21.5 | -23.5 | -11.2 | 0.2 |
| Peru 1991-92 | Latin America and Carib. | 2 | 1 | 0 | -2.3 | -1.6 | 25.1 | 151.0 | -6.4 | 4.5 |
| Peru 1996 | Latin America and Carib. | 3 | 1 | 4 | -2.3 | -1.6 | 19.4 | 45.5 | -21.2 | 6.1 |
| Peru 2000 | Latin America and Carib. | 4 | 1 | 2 | 1.0 | -2.9 | 19.6 | 86.6 | -18.9 | 5.3 |
| Peru Continuous DHS 2004-06 | Latin America and Carib. | 5 | 0 | 0 | -1.2 | | 34.0 | 146.0 | 4.8 | -6.2 |

**(Continued...)**

**Table A1. – *Continued***

| Survey | Region | Phase of DHS | Percentages of: Births with incomplete birthdate | Percentages of: Deaths with flagged age at death | Deviation for sex ratio at birth | Deviation for sex ratio of neonatal deaths | Deviation of percent of infant deaths that are neonatal | Relative deviation (x100) for deaths at 12 months | Relative deviation (x100) for deaths at 7 days | Relative deviation (x100) for births in year before boundary |
|---|---|---|---|---|---|---|---|---|---|---|
| Peru Continuous DHS 2007-08 | Latin America and Carib. | 5 | 0 | 0 | -2.3 | | 34.3 | 133.1 | -72.8 | -3.9 |
| Peru Continuous DHS 2010 | Latin America and Carib. | 6 | 0 | 0 | -6.3 | | 15.3 | 15.4 | 33.5 | -0.2 |
| Peru Continuous DHS 2011 | Latin America and Carib. | 6 | 0 | 0 | -0.3 | | -20.5 | 37.9 | 7.7 | 3.4 |
| Philippines 1993 | South and SE Asia | 3 | 1 | 0 | 2.4 | -7.5 | 19.6 | 45.0 | 28.6 | 2.2 |
| Philippines 1998 | South and SE Asia | 3 | 1 | 1 | 2.4 | -3.1 | 25.0 | 239.1 | 70.3 | 2.5 |
| Philippines 2003 | South and SE Asia | 4 | 0 | 2 | 3.9 | | -18.9 | 155.2 | 15.4 | 6.2 |
| Philippines 2008 | South and SE Asia | 5 | 0 | 1 | 6.5 | | -14.9 | 170.6 | 12.4 | 2.3 |
| Rwanda 1992 | Eastern Africa | 2 | 5 | 3 | -2.3 | -2.9 | 39.4 | 148.0 | -7.5 | 2.3 |
| Rwanda 2000 | Eastern Africa | 4 | 3 | 8 | -1.5 | -0.5 | 13.6 | 99.9 | -33.5 | 12.8 |
| Rwanda 2005 | Eastern Africa | 5 | 2 | 1 | 1.4 | -3.2 | 34.8 | 154.0 | -31.6 | 0.9 |
| Rwanda 2010 | Eastern Africa | 6 | 0 | 0 | 0.7 | -2.7 | 29.1 | 128.5 | -33.4 | 2.5 |
| Sao Tome and Principe 2008-09 | Middle Africa | 5 | 1 | 5 | 0.1 | -4.3 | -33.0 | 265.1 | 51.3 | 11.4 |
| Senegal 1992-93 | Western Africa | 2 | 33 | 1 | -4.1 | 0.6 | 11.2 | 131.0 | 5.3 | 11.1 |
| Senegal 1997 | Western Africa | 3 | 12 | 1 | 0.3 | -0.8 | 22.4 | 181.0 | -5.9 | 4.4 |
| Senegal 2005 | Western Africa | 4 | 8 | 13 | 1.6 | -1.1 | 8.1 | 54.4 | -21.9 | 7.3 |
| Senegal 2010-11 | Western Africa | 6 | 14 | 18 | 1.9 | -1.0 | -28.0 | 99.4 | -11.2 | 9.2 |
| Senegal Continuous DHS 2012-13 | Western Africa | 6 | 2 | 23 | 1.1 | -0.4 | -6.6 | 17.0 | 26.8 | 6.9 |
| Sierra Leone 2008 | Western Africa | 5 | 2 | 7 | -3.1 | -2.8 | 31.9 | 208.4 | -24.4 | 26.3 |
| South Africa 1998 | Southern Africa | 3 | 2 | 13 | -1.9 | -7.2 | -15.8 | 205.7 | 21.3 | 10.9 |
| Swaziland 2006-07 | Southern Africa | 5 | 0 | 7 | 1.5 | -10.2 | 48.7 | 229.2 | 23.9 | 14.8 |
| Tanzania 1991-92 | Eastern Africa | 2 | 13 | 5 | -1.8 | -2.8 | 35.5 | 72.4 | -22.5 | 7.0 |
| Tanzania 1996 | Eastern Africa | 3 | 6 | 9 | 2.5 | -2.3 | 20.5 | 115.9 | -32.7 | 10.8 |
| Tanzania 1999 | Eastern Africa | 4 | 3 | 2 | 2.0 | 0.2 | -20.9 | 166.9 | -64.7 | 2.4 |
| Tanzania 2004-05 | Eastern Africa | 4 | 1 | 4 | -0.5 | -0.2 | 11.1 | 147.9 | -48.4 | 2.8 |
| Tanzania 2010 | Eastern Africa | 6 | 0 | 4 | -3.3 | -1.3 | 26.8 | 51.9 | -38.8 | 4.2 |
| Timor-Leste 2009-10 | South and SE Asia | 6 | 0 | 9 | 0.3 | -2.4 | 28.3 | 111.5 | -19.3 | 14.2 |
| Togo 1998 | Western Africa | 3 | 24 | 8 | -1.5 | -1.1 | 8.5 | -35.5 | 13.6 | 7.9 |
| Turkey 1993 | N Africa / W Asia / Europe | 3 | 1 | 2 | 1.5 | -2.9 | 20.5 | 152.2 | 16.0 | 14.3 |
| Turkey 1998 | N Africa / W Asia / Europe | 4 | 6 | 5 | 2.0 | 0.2 | 6.0 | 190.2 | 20.8 | 9.0 |
| Turkey 2003 | N Africa / W Asia / Europe | 4 | 3 | 2 | -0.9 | -3.5 | 36.0 | 100.6 | -31.6 | 7.4 |

**(Continued...)**

**Table A1. – *Continued***

| Survey | Region | Phase of DHS | Percentages of: Births with incomplete birthdate | Percentages of: Deaths with flagged age at death | Deviation for sex ratio at birth | Deviation for sex ratio of neonatal deaths | Deviation of percent of infant deaths that are neonatal | Relative deviation (x100) for deaths at 12 months | Relative deviation (x100) for deaths at 7 days | Relative deviation (x100) for births in year before boundary |
|---|---|---|---|---|---|---|---|---|---|---|
| Uganda 1995 | Eastern Africa | 3 | 5 | 6 | -7.2 | -4.9 | 58.1 | 63.7 | -54.1 | 16.0 |
| Uganda 2000-01 | Eastern Africa | 4 | 3 | 8 | -1.2 | -4.3 | 34.4 | 35.7 | -32.7 | 12.3 |
| Uganda 2006 | Eastern Africa | 5 | 2 | 3 | -0.8 | -2.4 | -36.0 | 146.9 | -37.0 | 12.5 |
| Uganda 2011 | Eastern Africa | 6 | 1 | 2 | -2.4 | -0.7 | 22.7 | 176.7 | -56.0 | 3.1 |
| Ukraine 2007 | N Africa / W Asia / Europe | 5 | 0 | 12 | 8.4 | | -142.8 | | 157.6 | -2.0 |
| Uzbekistan 1996 | Central Asia | 3 | 0 | 1 | -2.2 | -3.4 | 21.0 | 173.8 | 123.4 | -6.0 |
| Vietnam 1997 | South and SE Asia | 3 | 0 | 0 | 4.3 | 0.1 | -40.9 | 225.6 | 51.5 | 13.1 |
| Vietnam 2002 | South and SE Asia | 4 | 1 | 0 | -1.0 | | 45.8 | 81.4 | 34.9 | 3.4 |
| Yemen 1991-92 | N Africa / W Asia / Europe | 2 | 46 | 9 | -0.3 | 0.8 | 11.3 | 293.5 | -43.0 | 17.4 |
| Zambia 1992 | Eastern Africa | 2 | 2 | 2 | -5.1 | 0.0 | 15.9 | 14.1 | -3.1 | 5.0 |
| Zambia 1996 | Eastern Africa | 3 | 1 | 7 | -5.8 | -2.0 | 12.5 | 75.9 | -24.2 | 10.5 |
| Zambia 2001-02 | Eastern Africa | 4 | 1 | 7 | -0.2 | -5.6 | 39.8 | 88.4 | -53.9 | 8.0 |
| Zambia 2007 | Eastern Africa | 5 | 1 | 9 | -6.6 | -1.1 | 25.4 | 100.8 | -39.9 | -0.3 |
| Zimbabwe 1994 | Eastern Africa | 3 | 0 | 4 | -2.4 | -3.9 | 28.1 | 73.4 | -31.5 | 1.6 |
| Zimbabwe 1999 | Eastern Africa | 4 | 0 | 13 | 2.1 | -3.8 | 25.9 | 68.4 | -50.8 | 1.9 |
| Zimbabwe 2005-06 | Eastern Africa | 5 | 0 | 10 | 2.9 | -6.8 | 21.5 | 196.7 | -42.4 | 1.9 |
| Zimbabwe 2010-11 | Eastern Africa | 6 | 1 | 0 | -2.7 | -1.0 | -1.1 | -26.9 | -46.8 | 7.2 |

**Table A2. Numbers of observed and estimated births and under-5 deaths, and estimated percentages of births and under-5 deaths that were omitted or displaced, all 192 surveys and births in the past 10 years**

| Survey | Region | Phase of DHS | Observed births | Estimated births | Estimated % of births | | Observed U5 deaths | Estimated U5 deaths | Estimated % of deaths | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Omitted | Displaced | | | Omitted | Displaced |
| Albania 2008-09 | N Africa / W Asia / Europe | 5 | 4,126 | 4,233 | 3 | 1 | 88 | 90 | 3 | 3 |
| Armenia 2000 | N Africa / W Asia / Europe | 4 | 3,957 | 4,150 | 5 | 0 | 180 | 199 | 10 | 1 |
| Armenia 2005 | N Africa / W Asia / Europe | 5 | 3,074 | 3,310 | 8 | 1 | 96 | 107 | 11 | 6 |
| Armenia 2010 | N Africa / W Asia / Europe | 6 | 2,796 | 2,910 | 4 | 1 | 58 | 64 | 10 | 7 |
| Azerbaijan 2006 | N Africa / W Asia / Europe | 5 | 4,695 | 5,034 | 7 | 1 | 250 | 285 | 14 | 4 |
| Bangladesh 1993-94 | South and SE Asia | 3 | 15,438 | 15,641 | 1 | 1 | 2,066 | 2,148 | 4 | 4 |
| Bangladesh 1996-97 | South and SE Asia | 3 | 13,405 | 13,951 | 4 | 1 | 1,534 | 1,668 | 9 | 3 |
| Bangladesh 1999-00 | South and SE Asia | 4 | 13,865 | 14,028 | 1 | 1 | 1,367 | 1,439 | 5 | 3 |
| Bangladesh 2004 | South and SE Asia | 4 | 14,209 | 14,349 | 1 | 1 | 1,244 | 1,287 | 4 | 3 |
| Bangladesh 2007 | South and SE Asia | 5 | 12,798 | 13,135 | 3 | 1 | 866 | 961 | 11 | 2 |
| Bangladesh 2011 | South and SE Asia | 6 | 18,622 | 18,753 | 1 | 1 | 1,075 | 1,111 | 3 | 3 |
| Benin 1996 | Western Africa | 3 | 10,104 | 10,258 | 2 | 1 | 1,547 | 1,607 | 4 | 2 |
| Benin 2001 | Western Africa | 4 | 10,493 | 10,687 | 2 | 1 | 1,478 | 1,563 | 6 | 4 |
| Benin 2006 | Western Africa | 5 | 31,136 | 31,470 | 1 | 2 | 3,603 | 3,734 | 4 | 4 |
| Benin 2011-12 | Western Africa | 6 | 26,495 | 27,114 | 2 | 2 | 1,738 | 1,815 | 4 | 5 |
| Bolivia 1994 | Latin America and Carib. | 3 | 12,410 | 12,739 | 3 | 2 | 1,427 | 1,573 | 10 | 7 |
| Bolivia 1998 | Latin America and Carib. | 3 | 14,747 | 14,818 | 0 | 1 | 1,342 | 1,370 | 2 | 7 |
| Bolivia 2003 | Latin America and Carib. | 4 | 21,608 | 21,890 | 1 | 1 | 1,826 | 1,906 | 4 | 6 |
| Bolivia 2008 | Latin America and Carib. | 5 | 18,328 | 18,444 | 1 | 2 | 1,234 | 1,270 | 3 | 7 |
| Brazil 1991 | Latin America and Carib. | 2 | 6,793 | 7,004 | 3 | 1 | 695 | 755 | 9 | 3 |
| Brazil 1996 | Latin America and Carib. | 3 | 10,872 | 11,001 | 1 | 2 | 574 | 636 | 11 | 3 |
| Burkina Faso 1993 | Western Africa | 2 | 11,438 | 11,595 | 1 | 4 | 1,982 | 2,037 | 3 | 11 |
| Burkina Faso 1998-99 | Western Africa | 3 | 11,805 | 11,822 | 0 | 2 | 2,245 | 2,245 | 0 | 5 |
| Burkina Faso 2003 | Western Africa | 4 | 21,421 | 21,579 | 1 | 2 | 3,490 | 3,567 | 2 | 7 |
| Burkina Faso 2010 | Western Africa | 6 | 29,773 | 30,227 | 2 | 2 | 3,693 | 3,832 | 4 | 4 |
| Burundi 2010 | Eastern Africa | 6 | 14,051 | 14,159 | 1 | 1 | 1,455 | 1,485 | 2 | 4 |
| Cambodia 2000 | South and SE Asia | 4 | 20,338 | 20,645 | 2 | 4 | 2,302 | 2,409 | 5 | 7 |
| Cambodia 2005 | South and SE Asia | 5 | 17,643 | 17,910 | 2 | 1 | 1,744 | 1,769 | 1 | 5 |
| Cambodia 2010 | South and SE Asia | 6 | 16,198 | 16,334 | 1 | 1 | 1,011 | 1,046 | 3 | 4 |
| Cameroon 1991 | Middle Africa | 2 | 6,432 | 6,520 | 1 | 1 | 774 | 797 | 3 | 4 |
| Cameroon 1998 | Middle Africa | 3 | 7,956 | 8,201 | 3 | 1 | 1,038 | 1,081 | 4 | 3 |

**(Continued...)**

**Table A2. – *Continued***

| Survey | Region | Phase of DHS | Observed births | Estimated births | Estimated % of births | | Observed U5 deaths | Estimated U5 deaths | Estimated % of deaths | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Omitted | Displaced | | | Omitted | Displaced |
| Cameroon 2004 | Middle Africa | 4 | 15,374 | 15,496 | 1 | 2 | 1,967 | 1,988 | 1 | 4 |
| Cameroon 2011 | Middle Africa | 6 | 22,311 | 22,729 | 2 | 1 | 2,448 | 2,563 | 5 | 4 |
| Central African Republic 1994-95 | Middle Africa | 3 | 9,364 | 9,453 | 1 | 1 | 1,320 | 1,337 | 1 | 4 |
| Chad 1996-97 | Middle Africa | 3 | 14,306 | 14,537 | 2 | 2 | 2,427 | 2,467 | 2 | 6 |
| Chad 2004 | Middle Africa | 4 | 11,246 | 11,260 | 0 | 3 | 1,993 | 1,986 | 0 | 7 |
| Colombia 1990 | Latin America and Carib. | 2 | 7,572 | 7,981 | 5 | 0 | 243 | 295 | 21 | 7 |
| Colombia 1995 | Latin America and Carib. | 3 | 9,996 | 10,057 | 1 | 1 | 353 | 368 | 4 | 2 |
| Colombia 2000 | Latin America and Carib. | 4 | 9,463 | 9,480 | 0 | 0 | 252 | 255 | 1 | 3 |
| Colombia 2005 | Latin America and Carib. | 5 | 29,964 | 30,211 | 1 | 0 | 729 | 735 | 1 | 2 |
| Colombia 2010 | Latin America and Carib. | 6 | 36,945 | 37,316 | 1 | 0 | 768 | 782 | 2 | 3 |
| Comoros 1996 | Eastern Africa | 3 | 4,049 | 4,143 | 2 | 2 | 406 | 431 | 6 | 5 |
| Congo (Brazzaville) 2005 | Middle Africa | 5 | 8,848 | 8,952 | 1 | 1 | 962 | 1,012 | 5 | 3 |
| Congo (Brazzaville) 2011-12 | Middle Africa | 6 | 17,259 | 17,524 | 2 | 0 | 1,161 | 1,229 | 6 | 5 |
| Congo Democratic Republic 2007 | Middle Africa | 5 | 16,559 | 16,887 | 2 | 1 | 2,170 | 2,242 | 3 | 3 |
| Cote d'Ivoire 1994 | Western Africa | 3 | 13,778 | 13,875 | 1 | 2 | 1,829 | 1,850 | 1 | 6 |
| Cote d'Ivoire 1998-99 | Western Africa | 3 | 3,931 | 4,008 | 2 | 3 | 614 | 640 | 4 | 8 |
| Cote d'Ivoire 2011-12 | Western Africa | 6 | 14,993 | 15,202 | 1 | 2 | 1,524 | 1,595 | 5 | 5 |
| Dominican Republic 1991 | Latin America and Carib. | 2 | 7,948 | 8,030 | 1 | 1 | 444 | 466 | 5 | 4 |
| Dominican Republic 1996 | Latin America and Carib. | 3 | 9,102 | 9,212 | 1 | 0 | 515 | 538 | 4 | 2 |
| Dominican Republic 1999 | Latin America and Carib. | 4 | 1,260 | 1,370 | 9 | 1 | 55 | 70 | 28 | 2 |
| Dominican Republic 2002 | Latin America and Carib. | 4 | 23,133 | 23,268 | 1 | 1 | 906 | 925 | 2 | 2 |
| Dominican Republic 2007 | Latin America and Carib. | 5 | 23,511 | 23,853 | 1 | 1 | 827 | 888 | 7 | 2 |
| Egypt 1992 | N Africa / W Asia / Europe | 2 | 18,104 | 18,215 | 1 | 2 | 1,757 | 1,787 | 2 | 14 |
| Egypt 1995 | N Africa / W Asia / Europe | 3 | 25,432 | 25,816 | 2 | 1 | 2,242 | 2,336 | 4 | 10 |
| Egypt 2000 | N Africa / W Asia / Europe | 4 | 22,915 | 23,117 | 1 | 2 | 1,448 | 1,468 | 1 | 12 |
| Egypt 2005 | N Africa / W Asia / Europe | 5 | 27,163 | 27,213 | 0 | 1 | 1,236 | 1,246 | 1 | 6 |
| Egypt 2008 | N Africa / W Asia / Europe | 5 | 21,786 | 21,964 | 1 | 2 | 676 | 720 | 7 | 7 |
| Ethiopia 2000 | Eastern Africa | 4 | 21,755 | 22,192 | 2 | 1 | 3,493 | 3,607 | 3 | 5 |
| Ethiopia 2005 | Eastern Africa | 5 | 20,342 | 20,523 | 1 | 2 | 2,347 | 2,366 | 1 | 7 |
| Ethiopia 2011 | Eastern Africa | 6 | 23,780 | 24,078 | 1 | 1 | 2,274 | 2,292 | 1 | 5 |
| Gabon 2000 | Middle Africa | 4 | 8,466 | 8,685 | 3 | 1 | 699 | 723 | 3 | 3 |
| Gabon 2012 | Middle Africa | 6 | 11,227 | 11,621 | 4 | 1 | 623 | 692 | 11 | 7 |

**(Continued...)**

**Table A2. – *Continued***

| Survey | Region | Phase of DHS | Observed births | Estimated births | Estimated % of births | | Observed U5 deaths | Estimated U5 deaths | Estimated % of deaths | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Omitted | Displaced | | | Omitted | Displaced |
| Ghana 1993 | Western Africa | 3 | 7,235 | 7,327 | 1 | 4 | 823 | 864 | 5 | 11 |
| Ghana 1998 | Western Africa | 4 | 6,611 | 6,641 | 0 | 1 | 628 | 642 | 2 | 5 |
| Ghana 2003 | Western Africa | 4 | 7,467 | 7,614 | 2 | 2 | 718 | 758 | 5 | 4 |
| Ghana 2008 | Western Africa | 5 | 5,881 | 5,994 | 2 | 2 | 432 | 456 | 6 | 6 |
| Guatemala 1995 | Latin America and Carib. | 3 | 19,384 | 19,541 | 1 | 1 | 1,357 | 1,397 | 3 | 5 |
| Guinea 1999 | Western Africa | 4 | 12,177 | 12,419 | 2 | 2 | 2,043 | 2,104 | 3 | 6 |
| Guinea 2005 | Western Africa | 5 | 13,539 | 13,805 | 2 | 3 | 2,210 | 2,248 | 2 | 7 |
| Guinea 2012 | Western Africa | 6 | 14,142 | 14,309 | 1 | 2 | 1,620 | 1,648 | 2 | 9 |
| Guyana 2009 | Latin America and Carib. | 5 | 4,564 | 4,652 | 2 | 1 | 171 | 190 | 11 | 3 |
| Haiti 1994-95 | Latin America and Carib. | 3 | 6,855 | 6,982 | 2 | 1 | 861 | 890 | 3 | 3 |
| Haiti 2000 | Latin America and Carib. | 4 | 13,426 | 14,133 | 5 | 1 | 1,590 | 1,728 | 9 | 3 |
| Haiti 2005-06 | Latin America and Carib. | 5 | 12,008 | 12,520 | 4 | 1 | 1,089 | 1,180 | 8 | 3 |
| Haiti 2012 | Latin America and Carib. | 6 | 14,001 | 14,146 | 1 | 1 | 1,158 | 1,221 | 5 | 3 |
| Honduras 2005-06 | Latin America and Carib. | 5 | 23,149 | 23,346 | 1 | 1 | 785 | 826 | 5 | 2 |
| Honduras 2011-12 | Latin America and Carib. | 6 | 21,081 | 21,603 | 2 | 0 | 563 | 661 | 17 | 1 |
| India 1992-93 | South and SE Asia | 2 | 128,372 | 130,574 | 2 | 1 | 13,764 | 14,678 | 7 | 4 |
| India 1998-99 | South and SE Asia | 4 | 120,082 | 122,912 | 2 | 2 | 11,150 | 12,125 | 9 | 4 |
| India 2005-06 | South and SE Asia | 5 | 109,384 | 112,026 | 2 | 1 | 8,553 | 9,422 | 10 | 2 |
| Indonesia 1991 | South and SE Asia | 2 | 33,983 | 34,987 | 3 | 2 | 3,263 | 3,533 | 8 | 8 |
| Indonesia 1994 | South and SE Asia | 3 | 38,627 | 38,977 | 1 | 1 | 3,247 | 3,295 | 1 | 8 |
| Indonesia 1997 | South and SE Asia | 3 | 36,077 | 36,718 | 2 | 2 | 2,248 | 2,383 | 6 | 8 |
| Indonesia 2002-03 | South and SE Asia | 4 | 32,706 | 33,056 | 1 | 2 | 1,627 | 1,758 | 8 | 4 |
| Indonesia 2007 | South and SE Asia | 5 | 37,578 | 38,354 | 2 | 1 | 1,740 | 1,855 | 7 | 8 |
| Indonesia 2012 | South and SE Asia | 6 | 36,792 | 37,069 | 1 | 1 | 1,430 | 1,481 | 4 | 3 |
| Jordan 1990 | N Africa / W Asia / Europe | 2 | 16,562 | 16,786 | 1 | 1 | 666 | 757 | 14 | 7 |
| Jordan 1997 | N Africa / W Asia / Europe | 3 | 12,415 | 12,517 | 1 | 1 | 397 | 403 | 1 | 6 |
| Jordan 2002 | N Africa / W Asia / Europe | 4 | 12,645 | 12,826 | 1 | 1 | 354 | 384 | 9 | 6 |
| Jordan 2007 | N Africa / W Asia / Europe | 5 | 20,464 | 20,561 | 0 | 0 | 431 | 457 | 6 | 5 |
| Jordan 2012 | N Africa / W Asia / Europe | 6 | 20,360 | 20,547 | 1 | 1 | 400 | 445 | 11 | 5 |
| Kazakhstan 1995 | Central Asia | 3 | 3,117 | 3,173 | 2 | 1 | 135 | 157 | 16 | 3 |
| Kazakhstan 1999 | Central Asia | 4 | 3,310 | 3,346 | 1 | 2 | 204 | 208 | 2 | 4 |
| Kenya 1993 | Eastern Africa | 3 | 12,460 | 12,921 | 4 | 1 | 1,053 | 1,149 | 9 | 6 |

**(Continued...)**

**Table A2. – *Continued***

| Survey | Region | Phase of DHS | Observed births | Estimated births | Estimated % of births | | Observed U5 deaths | Estimated U5 deaths | Estimated % of deaths | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Omitted | Displaced | | | Omitted | Displaced |
| Kenya 1998 | Eastern Africa | 3 | 11,283 | 11,453 | 2 | 1 | 1,072 | 1,163 | 8 | 4 |
| Kenya 2003 | Eastern Africa | 4 | 11,112 | 11,200 | 1 | 1 | 1,105 | 1,142 | 3 | 4 |
| Kenya 2008-09 | Eastern Africa | 5 | 11,645 | 11,817 | 1 | 2 | 877 | 900 | 3 | 6 |
| Kyrgyz Republic 1997 | Central Asia | 3 | 4,070 | 4,117 | 1 | 1 | 292 | 311 | 6 | 4 |
| Kyrgyz Republic 2012 | Central Asia | 6 | 7,686 | 7,879 | 3 | 1 | 229 | 272 | 19 | 3 |
| Lesotho 2004 | Southern Africa | 4 | 6,979 | 7,069 | 1 | 1 | 652 | 696 | 7 | 2 |
| Lesotho 2009 | Southern Africa | 6 | 7,250 | 7,343 | 1 | 1 | 705 | 729 | 3 | 4 |
| Liberia 2007 | Western Africa | 5 | 11,207 | 11,291 | 1 | 3 | 1,340 | 1,346 | 0 | 6 |
| Madagascar 1992 | Eastern Africa | 2 | 10,009 | 10,221 | 2 | 2 | 1,517 | 1,583 | 4 | 9 |
| Madagascar 1997 | Eastern Africa | 3 | 11,476 | 11,524 | 0 | 2 | 1,609 | 1,616 | 0 | 7 |
| Madagascar 2003-04 | Eastern Africa | 4 | 10,790 | 10,919 | 1 | 3 | 1,033 | 1,052 | 2 | 7 |
| Madagascar 2008-09 | Eastern Africa | 5 | 25,344 | 25,637 | 1 | 2 | 1,797 | 1,922 | 7 | 4 |
| Malawi 1992 | Eastern Africa | 2 | 8,759 | 8,955 | 2 | 2 | 1,802 | 1,913 | 6 | 5 |
| Malawi 2000 | Eastern Africa | 4 | 22,064 | 22,654 | 3 | 3 | 3,711 | 3,891 | 5 | 8 |
| Malawi 2004 | Eastern Africa | 4 | 19,969 | 20,289 | 2 | 2 | 2,622 | 2,711 | 3 | 5 |
| Malawi 2010 | Eastern Africa | 6 | 38,945 | 39,666 | 2 | 1 | 4,125 | 4,199 | 2 | 5 |
| Maldives 2009 | South and SE Asia | 5 | 7,234 | 7,305 | 1 | 0 | 177 | 188 | 6 | 2 |
| Mali 1995-96 | Western Africa | 3 | 20,426 | 20,894 | 2 | 2 | 4,375 | 4,545 | 4 | 4 |
| Mali 2001 | Western Africa | 4 | 26,138 | 26,247 | 0 | 2 | 5,348 | 5,365 | 0 | 8 |
| Mali 2006 | Western Africa | 5 | 28,201 | 28,483 | 1 | 3 | 5,102 | 5,143 | 1 | 7 |
| Moldova 2005 | N Africa / W Asia / Europe | 5 | 3,114 | 3,130 | 0 | 1 | 72 | 72 | 0 | 4 |
| Morocco 1992 | N Africa / W Asia / Europe | 2 | 10,598 | 10,645 | 0 | 1 | 802 | 824 | 3 | 4 |
| Morocco 2003-04 | N Africa / W Asia / Europe | 4 | 12,980 | 13,211 | 2 | 1 | 654 | 666 | 2 | 4 |
| Mozambique 1997 | Eastern Africa | 3 | 13,568 | 13,922 | 3 | 4 | 2,641 | 2,758 | 4 | 9 |
| Mozambique 2003 | Eastern Africa | 4 | 19,785 | 20,822 | 5 | 2 | 3,088 | 3,428 | 11 | 8 |
| Mozambique 2011 | Eastern Africa | 6 | 20,690 | 20,979 | 1 | 1 | 1,915 | 1,984 | 4 | 7 |
| Namibia 1992 | Southern Africa | 2 | 7,261 | 7,564 | 4 | 1 | 584 | 639 | 10 | 7 |
| Namibia 2000 | Southern Africa | 4 | 7,803 | 7,975 | 2 | 1 | 416 | 443 | 6 | 5 |
| Namibia 2006-07 | Southern Africa | 5 | 9,929 | 9,985 | 1 | 2 | 625 | 640 | 2 | 9 |
| Nepal 1996 | South and SE Asia | 3 | 14,572 | 14,709 | 1 | 0 | 1,753 | 1,803 | 3 | 2 |
| Nepal 2001 | South and SE Asia | 4 | 14,125 | 14,606 | 3 | 0 | 1,361 | 1,470 | 8 | 1 |
| Nepal 2006 | South and SE Asia | 5 | 11,991 | 12,355 | 3 | 0 | 859 | 962 | 12 | 2 |

**(Continued...)**

**Table A2. – *Continued***

| Survey | Region | Phase of DHS | Observed births | Estimated births | Estimated % of births | | Observed U5 deaths | Estimated U5 deaths | Estimated % of deaths | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Omitted | Displaced | | | Omitted | Displaced |
| Nepal 2011 | South and SE Asia | 6 | 11,292 | 11,548 | 2 | 1 | 658 | 729 | 11 | 2 |
| Nicaragua 1998 | Latin America and Carib. | 3 | 17,196 | 17,838 | 4 | 1 | 889 | 1,038 | 17 | 3 |
| Nicaragua 2001 | Latin America and Carib. | 4 | 15,043 | 15,277 | 2 | 2 | 620 | 740 | 19 | 3 |
| Niger 1992 | Western Africa | 2 | 13,539 | 13,977 | 3 | 2 | 3,707 | 3,930 | 6 | 4 |
| Niger 1998 | Western Africa | 3 | 15,483 | 15,599 | 1 | 2 | 3,774 | 3,818 | 1 | 3 |
| Niger 2006 | Western Africa | 5 | 18,798 | 19,099 | 2 | 5 | 3,413 | 3,514 | 3 | 15 |
| Niger 2012 | Western Africa | 6 | 25,355 | 25,395 | 0 | 4 | 3,116 | 3,118 | 0 | 12 |
| Nigeria 1990 | Western Africa | 2 | 15,912 | 16,096 | 1 | 3 | 2,632 | 2,701 | 3 | 7 |
| Nigeria 1999 | Western Africa | 4 | 12,524 | 12,719 | 2 | 2 | 1,462 | 1,524 | 4 | 3 |
| Nigeria 2003 | Western Africa | 4 | 11,573 | 11,728 | 1 | 0 | 2,091 | 2,158 | 3 | 1 |
| Nigeria 2008 | Western Africa | 5 | 55,594 | 56,287 | 1 | 2 | 8,087 | 8,352 | 3 | 3 |
| Pakistan 1990-91 | South and SE Asia | 2 | 14,900 | 15,016 | 1 | 5 | 1,677 | 1,705 | 2 | 7 |
| Pakistan 2006-07 | South and SE Asia | 5 | 19,100 | 19,770 | 4 | 2 | 1,675 | 1,848 | 10 | 5 |
| Paraguay 1990 | Latin America and Carib. | 2 | 8,356 | 8,448 | 1 | 0 | 354 | 378 | 7 | 1 |
| Peru 1991-92 | Latin America and Carib. | 2 | 18,726 | 19,052 | 2 | 1 | 1,541 | 1,631 | 6 | 6 |
| Peru 1996 | Latin America and Carib. | 3 | 35,043 | 35,575 | 2 | 1 | 2,171 | 2,270 | 5 | 4 |
| Peru 2000 | Latin America and Carib. | 4 | 29,225 | 29,524 | 1 | 1 | 1,545 | 1,638 | 6 | 5 |
| Peru Continuous DHS 2004-06 | Latin America and Carib. | 5 | 36,029 | 36,373 | 1 | 3 | 1,150 | 1,232 | 7 | 6 |
| Peru Continuous DHS 2007-08 | Latin America and Carib. | 5 | 20,445 | 20,735 | 1 | 1 | 569 | 609 | 7 | 5 |
| Peru Continuous DHS 2010 | Latin America and Carib. | 6 | 19,114 | 19,716 | 3 | 1 | 465 | 479 | 3 | 4 |
| Peru Continuous DHS 2011 | Latin America and Carib. | 6 | 18,876 | 18,918 | 0 | 1 | 454 | 461 | 2 | 3 |
| Philippines 1993 | South and SE Asia | 3 | 18,133 | 18,496 | 2 | 0 | 989 | 1,117 | 13 | 3 |
| Philippines 1998 | South and SE Asia | 3 | 16,279 | 16,558 | 2 | 1 | 760 | 844 | 11 | 8 |
| Philippines 2003 | South and SE Asia | 4 | 14,462 | 14,716 | 2 | 1 | 550 | 563 | 2 | 6 |
| Philippines 2008 | South and SE Asia | 5 | 13,389 | 13,791 | 3 | 0 | 460 | 480 | 4 | 6 |
| Rwanda 1992 | Eastern Africa | 2 | 11,098 | 11,414 | 3 | 1 | 1,544 | 1,666 | 8 | 5 |
| Rwanda 2000 | Eastern Africa | 4 | 14,872 | 15,031 | 1 | 3 | 2,689 | 2,733 | 2 | 6 |
| Rwanda 2005 | Eastern Africa | 5 | 16,669 | 17,057 | 2 | 2 | 2,523 | 2,701 | 7 | 7 |
| Rwanda 2010 | Eastern Africa | 6 | 17,587 | 17,802 | 1 | 1 | 1,464 | 1,564 | 7 | 7 |
| Sao Tome and Principe 2008-09 | Middle Africa | 5 | 3,702 | 3,735 | 1 | 2 | 226 | 243 | 8 | 7 |
| Senegal 1992-93 | Western Africa | 2 | 11,139 | 11,389 | 2 | 2 | 1,435 | 1,470 | 2 | 5 |
| Senegal 1997 | Western Africa | 3 | 14,677 | 14,815 | 1 | 1 | 1,757 | 1,826 | 4 | 4 |

**(Continued...)**

**Table A2. – *Continued***

| Survey | Region | Phase of DHS | Observed births | Estimated births | Estimated % of births | | Observed U5 deaths | Estimated U5 deaths | Estimated % of deaths | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Omitted | Displaced | | | Omitted | Displaced |
| Senegal 2005 | Western Africa | 4 | 21,042 | 21,300 | 1 | 1 | 2,408 | 2,473 | 3 | 2 |
| Senegal 2010-11 | Western Africa | 6 | 23,414 | 23,716 | 1 | 1 | 1,684 | 1,742 | 3 | 4 |
| Senegal Continuous DHS 2012-13 | Western Africa | 6 | 12,839 | 12,923 | 1 | 1 | 796 | 804 | 1 | 4 |
| Sierra Leone 2008 | Western Africa | 5 | 11,556 | 11,956 | 3 | 4 | 1,695 | 1,821 | 7 | 9 |
| South Africa 1998 | Southern Africa | 3 | 10,492 | 10,671 | 2 | 2 | 547 | 635 | 16 | 6 |
| Swaziland 2006-07 | Southern Africa | 5 | 5,373 | 5,520 | 3 | 2 | 529 | 611 | 16 | 6 |
| Tanzania 1991-92 | Eastern Africa | 2 | 15,240 | 15,506 | 2 | 1 | 2,018 | 2,117 | 5 | 5 |
| Tanzania 1996 | Eastern Africa | 3 | 13,021 | 13,298 | 2 | 2 | 1,656 | 1,732 | 5 | 6 |
| Tanzania 1999 | Eastern Africa | 4 | 6,198 | 6,264 | 1 | 1 | 868 | 884 | 2 | 7 |
| Tanzania 2004-05 | Eastern Africa | 4 | 16,040 | 16,133 | 1 | 1 | 1,809 | 1,841 | 2 | 5 |
| Tanzania 2010 | Eastern Africa | 6 | 15,219 | 15,488 | 2 | 1 | 1,212 | 1,261 | 4 | 3 |
| Timor-Leste 2009-10 | South and SE Asia | 6 | 19,695 | 19,862 | 1 | 2 | 1,367 | 1,435 | 5 | 5 |
| Togo 1998 | Western Africa | 3 | 14,153 | 14,309 | 1 | 1 | 1,750 | 1,780 | 2 | 3 |
| Turkey 1993 | N Africa / W Asia / Europe | 3 | 8,146 | 8,287 | 2 | 2 | 604 | 656 | 9 | 6 |
| Turkey 1998 | N Africa / W Asia / Europe | 4 | 7,303 | 7,381 | 1 | 1 | 400 | 408 | 2 | 6 |
| Turkey 2003 | N Africa / W Asia / Europe | 4 | 9,327 | 9,421 | 1 | 1 | 401 | 435 | 9 | 4 |
| Uganda 1995 | Eastern Africa | 3 | 13,157 | 13,892 | 6 | 2 | 1,746 | 1,943 | 11 | 4 |
| Uganda 2000-01 | Eastern Africa | 4 | 13,272 | 13,540 | 2 | 2 | 1,768 | 1,878 | 6 | 4 |
| Uganda 2006 | Eastern Africa | 5 | 16,125 | 16,330 | 1 | 2 | 1,990 | 2,086 | 5 | 7 |
| Uganda 2011 | Eastern Africa | 6 | 15,180 | 15,438 | 2 | 1 | 1,368 | 1,430 | 5 | 7 |
| Ukraine 2007 | N Africa / W Asia / Europe | 5 | 2,552 | 2,659 | 4 | 1 | 47 | 54 | 15 | 2 |
| Uzbekistan 1996 | Central Asia | 3 | 4,752 | 4,832 | 2 | 1 | 244 | 267 | 10 | 5 |
| Vietnam 1997 | South and SE Asia | 3 | 7,135 | 7,271 | 2 | 2 | 294 | 304 | 3 | 5 |
| Vietnam 2002 | South and SE Asia | 4 | 5,327 | 5,399 | 1 | 1 | 155 | 181 | 16 | 2 |
| Yemen 1991-92 | N Africa / W Asia / Europe | 2 | 15,911 | 16,013 | 1 | 3 | 1,974 | 2,009 | 2 | 10 |
| Zambia 1992 | Eastern Africa | 2 | 11,860 | 12,184 | 3 | 1 | 1,841 | 1,905 | 3 | 4 |
| Zambia 1996 | Eastern Africa | 3 | 13,253 | 13,678 | 3 | 2 | 2,203 | 2,317 | 5 | 5 |
| Zambia 2001-02 | Eastern Africa | 4 | 12,933 | 13,199 | 2 | 2 | 1,867 | 2,011 | 8 | 6 |
| Zambia 2007 | Eastern Africa | 5 | 11,845 | 12,310 | 4 | 2 | 1,365 | 1,447 | 6 | 5 |
| Zimbabwe 1994 | Eastern Africa | 3 | 8,246 | 8,434 | 2 | 0 | 557 | 608 | 9 | 4 |
| Zimbabwe 1999 | Eastern Africa | 4 | 6,978 | 7,102 | 2 | 1 | 582 | 620 | 7 | 4 |
| Zimbabwe 2005-06 | Eastern Africa | 5 | 9,875 | 10,107 | 2 | 1 | 623 | 708 | 14 | 5 |
| Zimbabwe 2010-11 | Eastern Africa | 6 | 10,075 | 10,226 | 1 | 1 | 704 | 725 | 3 | 3 |