USAID
FROM THE AMERICAN PEOPLE

METHODS TO ESTIMATE UNDER-15 FERTILITY USING DEMOGRAPHIC AND HEALTH SURVEYS DATA

DHS METHODOLOGICAL REPORTS 23

SEPTEMBER 2018

This publication was produced for review by the United States Agency for International Development. It was prepared by Thomas W. Pullum, Trevor Croft, and Kerry L. D. MacQuarrie.

DHS Methodological Reports No. 23

Methods to Estimate Under-15 Fertility Using Demographic and Health Surveys Data

Thomas W. Pullum[1]
Trevor Croft[1]
Kerry L. D. MacQuarrie[2]

ICF
Rockville, Maryland, USA

September 2018

[1] The DHS Program, ICF
[2] The DHS Program, Avenir Health

Corresponding author: Thomas W. Pullum, International Health and Development, ICF, 530 Gaither Road, Suite 500, Rockville, MD 20850, USA; telephone: +1 301-572-0950; fax: +1 301-572-0999; email: tom.pullum@icf.com

Acknowledgments: Helpful comments from Fred Arnold and Sunita Kishor are gratefully acknowledged.

Editor: Bryant Robey
Document Production: Joan Wardell

This study was carried out with support provided by the United States Agency for International Development (USAID) through The DHS Program (#AID-OAA-C-13-00095). 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. Additional information about The DHS Program can be obtained from ICF, 530 Gaither Road, Suite 500, Rockville, MD 20850 USA; telephone: +1 301-572-0200, fax: +1 301-572-0999, email: info@DHSprogram.com, internet: www.DHSprogram.com.

Recommended citation:

Pullum, Thomas W., Trevor Croft, and Kerry L. D. MacQuarrie. 2018. *Methods to Estimate Under-15 Fertility Using Demographic and Health Surveys Data*. DHS Methodological Reports No. 23. Rockville, Maryland, USA: ICF.

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 methods used will enhance the accuracy and depth of information collected by The DHS Program and relied on by policymakers and program managers in low- and middle-income countries.

While data quality is a main topic 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 U.S. 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

The DHS Program

ABSTRACT

Early childbearing carries serious risks to the health of both the child and the mother. International guidelines classify births before age 18 as high-risk births, and births before age 15 are of even greater concern. Early marriage and pregnancy are also interpreted as negative indicators of child protection, and may severely limit an adolescent girl's educational opportunities.

DHS surveys are a major source of fertility estimates for age 15-49, the age range of eligibility for the women's interview and the collection of birth histories. Until recently, DHS surveys were not used for the calculation of fertility rates below age 15. This methodological report provides technical details for calculating fertility rates for age 12-14 and age 10-14 during the 3 years before the survey, and age 10-14 during the 5 years before the survey, the standard time intervals for DHS age-specific fertility rates. The under-15 births, which go into the numerator of the fertility rates, are obtained from the birth histories of women age 15-19 at the time of the survey.

The central question is how to deal with the left censoring of under-15 exposure, which goes into the denominator of fertility rates, but is truncated because girls under age 15 are not included in the surveys. To deal with censoring, rates for single years of age 10, 11, 12, 13, and 14 are constructed. DHS does not normally construct single-year rates, but they are of special interest below age 15 and convenient for calculating 3-year and 5-year rates. Single-year rates are only slightly affected by censoring. For example, the mean age observed for age 14 in the past 3 years is only about 12 days higher than 14.5. Pooled rates for age 12-14 or 10-14 are constructed as weighted averages of the single-year rates. Three alternative weighting methods are considered. The first uses information about girls age 10-14, who are included in the household survey but not in the women's survey. The second uses weights derived from the geometry of a Lexis diagram. These estimates are easier to calculate and are the ones currently available on STATcompiler. The third method is simply the arithmetic average of the single-year rates. These estimates are the easiest to calculate but lack a demographic rationale. The second and third methods, when relevant, also assume that 60% of the total exposure to age 10-14 is to age 12-14.

This report applies the alternative approaches to 67 DHS surveys conducted between 2001 and 2016. The three estimates are virtually indistinguishable in almost all surveys. When they differ, the main reasons appear to be that the assumption of 60% is not valid and/or there are irregularities in the reported age distribution that are probably traceable to digit preference or age displacement across the age 15 boundary or potential age displacement related to having an early birth. The first method, which makes the most use of the available data, is most sensitive to data quality. The third method, which makes the least use of available data, is least sensitive to data quality.

The report describes simulations of the effect of reducing the minimum age for eligibility from 15 to 14, or to 13, in terms of the expected additional number of under-15 births and the expected improvement in the precision of the estimated rate. The gain from lowering the age range of eligibility below age 15 would be surprisingly small. Most under-15 fertility occurs at age 14, and the birth histories of women 15-19 provide nearly complete information on births at age 14.

KEY WORDS: Adolescent fertility, adolescent birth rates, under-15 fertility, fertility rate construction

1 INTRODUCTION

1.1 Background

One of the most important objectives of The Demographic and Health Surveys (DHS) Program, from its inception in 1984, has been the estimation of fertility levels, trends, and differentials. The age range of eligibility for DHS surveys of individual women has always been age 15-49. Age-specific and total fertility rates have always been produced for this age range. The purpose of this report is to document recent efforts by DHS to produce estimates of fertility before age 15, while retaining age 15 as the lower end of the age range for eligibility for the survey.

Definitions of demographic rates that preceded the DHS considered the reproductive ages to be 15-44 or 15-49 (see, for example, Shryock, Siegel, and Associates 1971). Little fertility occurs after age 44 in most populations, at least without medical interventions. DHS (and the preceding World Fertility Survey) adopted the more expansive definition of the upper end of the reproductive ages by including age 45-49. There were at least two motivations for including this extra 5-year age interval. First, in contexts where age is not accurately measured, some childbearing that actually occurs before age 45 is incorrectly reported as occurring after age 45, and summary measures such as the total fertility rate (TFR) are biased downwards if that fertility is omitted. A second reason for including age 45-49 is to reduce the effect of truncation or censoring on the estimates of late fertility during the time periods before the survey.

Turning to the younger ages, the age range to define childhood is generally 0-17, the last 3 years of which (age 15-17) extend into the reproductive age range. Many reproductive health programs promote postponement of marriage and childbearing beyond the 18th birthday. There is even stronger interest in avoiding childbearing before the lower end of the age range for eligibility, the 15th birthday. Early pregnancy and childbearing are problematic from a number of perspectives, including the health of both the mother and the child. Early pregnancy and childbearing, whether or not the young woman is married, are typically incompatible with staying in school and are often treated as negative indicators of child protection.

Despite the importance of measurements of early fertility, DHS Occasional Paper 9 (Way 2014) described several reasons for not extending the lower end of eligibility for the women's interview below age 15. Issues include whether it is possible to obtain informed consent from someone under age 15 (see also WHO 2018), whether questions about sexual activity, pregnancy, childbearing, and contraceptive use are appropriate below age 15, and whether the responses will be complete in contexts where early pregnancy and childbearing are negatively sanctioned.[1]

DHS Comparative Report 45 (MacQuarrie, Mallick, and Allen 2017) shows that it is possible to infer a great deal about the recent reproductive health of age group 10-14 from the data currently being collected from age group 15-19, because their retrospective birth histories and other retrospective information include events before the 15th birthday. That report included estimates of fertility during the past 5 years for single years of age in the full age range 10-19.

[1] In this report, females age 15 and above will usually be referred to as women, and those below age 15 as girls.

This report will describe in considerably more detail than was possible in Comparative Report 45 how it is possible to estimate fertility before age 15 with DHS surveys. The documentation is needed because it is not intuitive that surveys with a lower age boundary of 15 can be a source of data for ages below 15. We will show how to pool and interpret estimates for age intervals—how, for example, to calculate the estimates for age 10-14 in the 3 years and 5 years before the survey.

Chapter 2 is concerned mainly with demographic and computational issues, and includes a detailed example. Chapter 3 applies and compares alternative approaches using 67 DHS surveys, selected as the most recent survey from all countries that have had at least one DHS survey since 2000. It also describes how to simulate the effect of reducing the minimum age of eligibility to 14 or 13, in terms of the additional number of births that would be expected and the impact on confidence intervals for the rates. Conclusions are given in Chapter 4. Appendix tables give numerical values of the estimates.

The report is intended for a wide audience, ranging from readers who are mainly interested in the methodological and computational details, to readers who are policy-oriented but are looking for some assurance that DHS surveys, as currently designed, can indeed be useful for estimating under-15 fertility. For some readers the report has more detail than necessary, and much of Chapter 2 could be skipped.

1.2 Calculating Age-Specific Fertility Rates with DHS Birth Histories

Descriptions of the procedures to construct fertility rates from the birth histories in DHS surveys are given in the online Guide to DHS Statistics and in other analysis reports, such as Methodological Report 12 (Schoumaker 2014), Methodological Report 21 (Pullum, Assaf, and Staveteig 2017), and Analytical Study 58 (Pullum and Assaf 2016). One of the most detailed descriptions of the calculation of fertility rates is provided in Pullum (2004). The description here will be brief.

For each woman, the essential information in a DHS survey for the calculation of fertility rates is the date of interview, the date of her birth, and the dates of birth of all her children as provided in the birth history.[2] The woman contributes births to the numerators of age-specific fertility rates according to when her successive age intervals occurred, relative to a reference time interval such as the 3 years before the survey (months 1-36 before each woman's month of interview). Using just her date of interview and the date of her own birth, her contributions to the denominators of the age-specific fertility rates are calculated.[3] The rates are calculated by adding the individual-level contributions to the numerators and to the denominators, and then dividing appropriately. The rates may include a factor of 1000. The age-specific rate for age 15-19 for the 3 years before the survey, for example, can be interpreted as the number of births while age 15-19 per 1,000 woman-years of exposure to that age range in that interval of time. The seven standard age intervals are 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, and 45-49. In symbols, the general definition for a fertility rate, ignoring the possible factor of 1000, is simply $r=b/e$, where r is the rate, b is the aggregated

[2] The date of interview has always included day, month, and year. The woman's date of birth and the dates of birth of her children have always included month and year. Beginning with surveys conducted in 2016, DHS now also collects the day of birth for all children, but not the day of the woman's birth. All fertility rates in DHS reports and on STATcompiler use only the month and year of the interview, the woman's birth, and the children's births, even when day of birth has also been included. This report will also ignore any information about day.

[3] Each woman contributes, to the denominator, the amount of time (measured in years) when she was within a specified age interval and within a specified time interval. The denominators of the rates are aggregations of woman-years of exposure to risk, not numbers of women.

number of births within a specified interval of age and time, and *e* is aggregated woman-years of exposure to the risk of a birth within the same specified interval of age and time.

DHS main reports also include the general fertility rate (GFR), which is essentially an age-specific rate for the full age range 15-49, and the total fertility rate (TFR), which is the sum of the seven age-specific rates, omitting the factor of 1000 and multiplying by 5. The TFR can be interpreted as the expected number of children that a woman would have between her 15th and 50th birthdays, if she survived for that full 35-year interval and had children at the rates experienced by women at successive ages during the reference interval of time. The GFR and TFR have well-understood definitions[4] that DHS does not propose to extend below age 15. This methodological report is thus limited to age-specific fertility in the age range 10 to 14, within the 3 or 5 years (1-36 months or 1-60 months) before the survey.[5]

DHS fertility rates are normally included in the DHS main reports for two different reference periods. First, age-specific rates, the GFR, and the TFR, are calculated at the national level and for urban-rural residence for the 3 years before the survey. A second table gives the TFR for subnational breakdowns such as region and wealth quintile, also using the 3 years before the survey. A 3-year interval is used to produce estimates that are as up to date as possible. A third table typically shows trends in age-specific rates at the national level across the most recent surveys, with 5-year reference intervals. A 5-year interval is used in order to have more births and less sampling error, and also to mesh with the standard 5-year width of age intervals and the typically 5-year intervals between surveys in a country. On STATcompiler (www.STATcompiler.com), users have the option of downloading age-specific rates with 3-year or 5-year reference periods. DHS main reports also typically include a table that gives the percentage of women age 15-19 who have ever had a birth or are currently pregnant, within single years of age, region, wealth quintile, etc.

Although eligibility for the individual interview is limited to women age 15 and older, there is no lower age limit for the birth history. Respondents are asked about the month and year of each birth, beginning with the first one. Any births before the 15th birthday will contribute to the numerator of an under-15 fertility rate. Such births during the 3 or 5 years before the survey are contributed solely by women who are age 15-19 at the time of the survey.

To fit age 10-14 into the existing framework of 5-year age-specific fertility rates, we would calculate an age-specific fertility rate for age 10-14 for the 3 years before the survey and/or the 5 years before the survey.

The rates discussed in this report are as follows:

Fertility rates for single years of age

Three years before the survey: rates for single years of age 12, 13, and 14

Five years before the survey: rates for single years of age 10, 11, 12, 13, and 14

[4] DHS uses a definition of the GFR that includes all births, regardless of age, and a denominator restricted to age 15-44, rather than 15-49.

[5] As with all the rates that are currently calculated, births and exposure during the month of interview are ignored, because only a fraction of that month is observed.

Normally, DHS would never suggest the calculation of rates for single years of age, regardless of the time interval. Single-year rates are not provided in main reports or on STATcompiler. However, within the age ranges 10-14 or 10-19 there is interest in identifying the age at which fertility first becomes non-negligible. Single-year rates are also convenient for the pooling strategies that will be described.

Fertility rates for age intervals

R1: three years before the survey and age interval 12-14

R2: three years before the survey and age interval 10-14

R3: five years before the survey and age interval 10-14

The only women age 15 or older at the time of the survey who will contribute to an under-15 rate for the 3 years before the survey are women currently age 15-17. In the past 3 years those women spent at least some time at age 12-14, but no time at younger ages, so a 3-year rate for age 10-14 cannot include any births at age 10-11. R2 and R3 are included because the main reports and STATcompiler include rates for 5-year age intervals 15-19 through 45-49 for the 3 years and 5 years before the survey, and for many users the natural extension to ages below 15 would be for a 5-year age interval, age 10-14. Some assumptions are required to estimate R2 from R1.

Most of this report is concerned with ways to calculate R1, R2, and R3 from the rates for single years of age. All of the options assume the prior calculation of the single-year rates. The calculation of the single-year rates is completely consistent with the procedures to estimate the rates for the 5-year age intervals 15-19 through 45-49.[6] The next chapter will go into details for alternative strategies to estimate the pooled rates R1, R2, and R3. Some readers might wish to skip those details and proceed to Chapter 3.

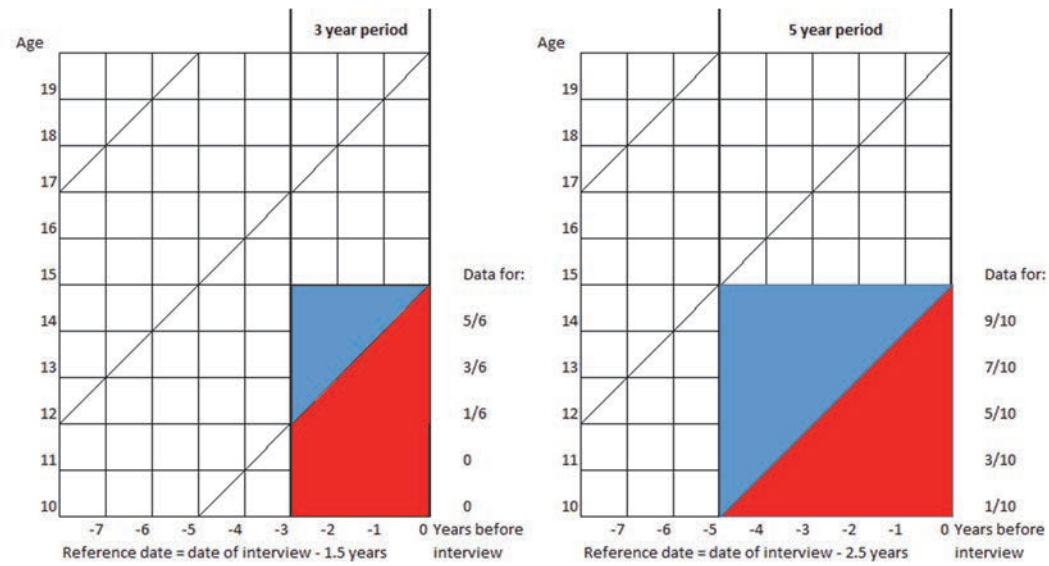
[6] The standard rates require the specifications that the initial month of age is 180 (age in months at the 15th birthday), the width of each age interval is 60 months (5 years), and there are seven age intervals (15-19 through 45-49). For the single-year rates for age 10 through 14, the specifications are that the initial month of age is 120 (age in months at the 10th birthday), the age interval is 12 months (one year), and there are five age intervals (10, 11, 12, 13, and 14). All other aspects of the calculations are identical.

2 METHODS

2.1 The Role of Censoring

With respect to data that are structured chronologically, such as birth histories, “censoring” refers to a pattern of truncation, relative to the underlying behavior that is being measured. Thinking graphically of time along a horizontal axis in which later events are to the right of earlier events, the reduction in exposure that results from omission of respondents younger than age 15 can be described as a type of left-censoring of the birth histories. It is not as serious as omission of births before age 15 would be, but it requires some adjustments during analysis.[7]

Figure 2.1 Lexis diagrams showing, in blue and red, the areas for which births and exposure are required for the fertility rates for age 10-14 during the 3 years and 5 years before the survey



A Lexis diagram is a graphical device to illustrate the structure of the data in terms of time and age, and can help to show how the data are incomplete. Much of the discussion in this chapter relates to the Lexis diagrams shown in Figure 2.1. The horizontal axis is time, expressed as years before the interview, and the vertical axis is years of age. “Exposure” refers to the combinations of age and time when the women or girls were alive. Horizontal and vertical lines within the figure mark off the boundaries of the years of age and time, respectively. Diagonal lines show the boundaries of 5-year age groups 10-14 and 15-19 (at the time of the survey) as they are backdated to earlier times and ages.

The colored rectangle on the left shows the area within which births and exposure are required to estimate a fertility rate for age 10-14 during the 3 years before the survey. The colored square on the right shows the area within which births and exposure are required to estimate a fertility rate for age 10-14 during the 5 years before the survey. In each blue triangle, births and exposure come from the data for women age 15-

[7] The general policy of DHS is not to adjust data. The procedures described in this report are minimal adaptations to the structure of the data.

19 at the time of the survey; the red area is the portion that will be estimated from the household survey or from assumptions of proportionality.

2.2 Bias within a Single Year of Age

In the Lexis diagram in the left of Figure 2.1, consider the exposure to age 14 during just the 12 months before the survey. The only exposure to this year of age, within this year of time, comes from women who are age 15 at the time of the survey. Assuming a uniform distribution of birth months across the year, about 1/12 of the women who are age 15 will have had one month of exposure to this combination of age and time; about 1/12 will have had two months of exposure; etc. One result of this pattern is that in the aggregate there is more exposure to the older half of age 14 than to the younger half. Another way to describe the pattern is that we are missing exposure to age 14 by girls who are still age 14 at the time of the survey and therefore are not eligible to provide a birth history.

A figure for the aggregated exposure to the combination of age 14 and the 12 months before the survey would have the shape of a right triangle. In such a triangle, the mean age observed would be two-thirds of the way from the 14th birthday to the 15th birthday. That is, on a continuous scale of age, the mean age for the available data on 14-year-olds is 14.67.

If we did not have left-censoring and could include, say, events and exposure to this combination of age and time from girls who were still age 14 at the time of the survey, then the mean age observed would be at the midpoint of the year of age, i.e., 14.50. Thus, the displacement toward later ages within age 14 amounts to 14.67-14.50=0.17 (one-sixth) of a year, or two months. Because the underlying probability of childbearing probably increases slightly with age, within age 14, the displacement tends to bias the estimate slightly in an upward direction. The "true" fertility rate for age 14 is slightly less than the estimate.

This kind of displacement, such that age in the observed data tends to be displaced upwards by two months relative to complete data, applies to all the squares in Figure 2.1 that represent a single year of age and a single year of time, and are bisected by a diagonal line. The single-year squares with full blue shading are not affected by displacement, so the bias within a single year of age becomes smaller as the time interval is extended backwards. For example, it can be shown that the mean age for age 14 during the 3 years before the survey is displaced upward by only 1/30 of a year (about 12 days). The 3-year estimate for age 13 is displaced upward by 1/18 of a year (about 20 days). The largest displacement for a 3-year estimate is for age 12, which is 1/6 of a year (two months). Among the 5-year pooled estimates, the displacement above the midpoint of the year of age is 1/54 of a year for age 14, 1/42 for age 13, 1/30 for age 12, 1/18 for age 11, and 1/6 for age 10.

The upward bias in the estimated rates due to this within-year age shift is probably negligible, especially if weighed against other potential sources of age misreporting. We suggest that the bias is in the preferred direction. For the policy-related uses to which the estimates will be put, a slight overestimate would be preferable to an underestimate. Any adjustment to compensate for this kind of displacement is outside the scope of this report and DHS policy.

2.3 Supplementing the Exposure to Age Intervals

The red area in the lower right of the Lexis diagrams represents data that are missing from the birth histories provided by interviewed women age 15 and above, but it is possible to supplement or augment the exposure, the denominators of the rates, by reference to the information in the household survey. Data collection for DHS surveys begins with visits to selected households and a listing of all occupants of those households. The members of all the households in the survey are the cases in the "PR" data file. The relevant information about each case includes sex; age (in years); whether a "usual" (de jure) resident; and whether "stayed in the household last night" (*de facto*) resident.

The household members who are included in the survey of women usually must satisfy only the following three criteria [8]: they are female; they are age 15-49; they are *de facto* residents of the household (i.e. they "stayed in the household last night"). Some surveys add a fourth requirement, that the women are ever-married. Surveys of ever-married women—EMW surveys—will be discussed later in the report. The eligible women who are successfully interviewed are the cases in the "IR" data file.

During data analysis it is possible to loosen the second criterion for inclusion in the survey, namely that age be in the range 15-49, by adding to the women's file those girls who are age 10-14 in the household survey and are excluded from the women's survey. These girls do not provide a birth history, but, to repeat, they provide information about exposure to the risk of early childbearing. We essentially supplement the women's file by adding to it girls age 10-14 who satisfy all the eligibility criteria other than age.

In the survey of women, month and year of birth are obtained directly from the women respondents. In the household survey, ages are provided for all members of the household by the "household respondent" and are given only as years of age at last birthday. However, a year and month of birth are required for all the standard calculations of fertility rates. For the girls age 10-14, two strategies are used to estimate the year and month of birth.

First, in all the surveys we have checked, the great majority of girls age 10-14 are co-residing with their mother. For these girls, their year and month of birth are included in the interviewed mother's birth history. There are only three circumstances in which a girl age 10-14 will not appear in the birth history of a woman in the household. First, the mother may have died; second, the mother may have survived but the daughter and mother are not living in the same household; third, the mother's current age may be 50 or above, in which case she too does not satisfy the age criterion and will not be interviewed.

If the year and month of birth of a girl age 10-14 are given in the co-resident mother's birth history, that information will take precedence over the years of age reported in the household survey. If the revised estimate of age is outside the range 10-14, then the case will not be added to the supplemented data. Likewise, if her revised age is under 10, then we omit her. If her revised age is 15 or above, then she will almost always have been included in the main survey of women; if she was not included (her age may have

[8] During a woman's interview, it can happen that her estimate of age is revised to be outside the range 15-49 and she is deemed ineligible to continue with the interview. Residence in a few surveys, such as the India surveys, is *de jure* rather than *de facto*.

been revised later or she may be a "refusal or nonresponse" case), then we omit her, because we do not want to change the data in the file of women in any way other than adding an age 10-14 category.

Second, for those girls age 10-14 who do not appear in a birth history, year and month of birth are estimated in a different way. For a girl age a (a=10, 11, 12, 13, or 14) it is assumed that her last birthday was six months before the interview. Thus, if doi is the date of interview, which is known, and dob is the date of birth, which is not known, and both are expressed with century month codes (cmc), then we estimate dob with dob=doi-12a-6.

This procedure for imputing a month and year of birth is similar to DHS procedures used in the construction of some standard DHS report tables. For example, a standard table on school attendance during the past year typically takes account of eligibility based on the child's age in some qualifying calendar month. Because children in the household survey have a year of age but do not have a year and month of birth, information about date of birth is borrowed from the birth histories when possible. For children who do not appear in the birth histories, DHS estimates date of birth randomly within a one-year range that is consistent with the stated age. Probabilistic imputation is not used in this report because different users following the same approach would not obtain exactly the same results. To avoid that type of uncertainty, to repeat, we estimate the child's cmc of birth with dob=doi-12a-6.

2.4 Calculating Fertility Rates for Pooled Ages: "PR" Method

Supplementing exposure as described above is relevant for the calculation of the pooled rates for ages 12-14 or 10-14. Within those intervals of age, there is a progressive censoring of cases in the earlier ages. Figure 2.1 clearly shows that there are far more observations of fertility at age 14, for example, than at age 12 or at age 10. If only the observed data for women age 15-19 at the time of the survey are used, this censoring has the effect of weighting the pooled rates toward the later ages within the age range. Because fertility increases within the age ranges 12-14 or 10-14, the pooled rate will be biased upwards.

Without any correction for this bias toward age 14, and using just the births and exposure from the woman's file, the pooled rate for age 12-14 during the past 3 years, for example, would be a weighted average of the single-year rates for ages 12, 13, and 14 in which the weights are the proportion of observed exposure at each age:

$$R1\_raw = \frac{(b_{12} + b_{13} + b_{14})}{(e_{12} + e_{13} + e_{14})} = \frac{(e_{12}r_{12} + e_{13}r_{13} + e_{14}r_{14})}{(e_{12} + e_{13} + e_{14})} = w_{12}r_{12} + w_{13}r_{13} + w_{14}r_{14}$$

where $w_{12} = e_{12}/(e_{12} + e_{13} + e_{14})$, etc.; here b is the number of births, e is the woman-years of exposure, and r=b/e is the rate for a single year of age. We emphasize that this formula for R1_raw includes a serious bias toward age 14 and is therefore too high. It is a starting point but should not be used.

When using the supplemented data from the household roster or PR file, the aggregated exposure to each year of age within the reference period of time is calculated. The totals will be represented with an uppercase E rather than a lowercase e. The single-year rates themselves are calculated exactly as before, completely from the births and exposure among women age 15-19 at the time of the survey, but the pooled rate is re-weighted[9] using E rather than e. Specifically, the re-weighted rate for age 12-14 will be

$$R1\_PR = W_{12}r_{12} + W_{13}r_{13} + W_{14}r_{14}$$

where $$W_{12} = E_{12}/(E_{12} + E_{13} + E_{14}) \text{ , etc.}$$

The re-weighted $R1$_PR is corrected for the bias toward age 14 that would otherwise occur in the censored data. The only remaining bias is the slight upward bias in each of the single-year rates because the observations for each year are shifted slightly past the midpoint of the year, by amounts given earlier, which we judge to be negligible and will not attempt to assess.

A different perspective on the function of the W weights is that they inflate the number of births and also the amount of exposure, within each single year of age, to the levels that would be expected in a survey with a lower age for eligibility of age 10 rather than 15.

The supplemented data can also be used to estimate R2_PR and R3_PR, the estimated rates for age 10-14 in the past 3 years and the past 5 years, respectively. Define E to be the total years of exposure to age 10-14 in the past 3 years and $P = (E_{12} + E_{13} + E_{14})/E$ to be the proportion of that total which was within age 12-14. Then under the assumption of no births to girls age 10-11, a necessary assumption because we have no observed births to age 10-11 in the past 3 years, the estimate of R2 will be

$$R2\_PR = P * R1\_PR \text{ .}$$

The estimate of R3_PR uses exposure to individual years of age 10, 11, 12, 13, and 14 during the past 5 years and the single-year rates for those ages during the past 5 years. Thus

$$R3\_PR = W_{10}r_{10} + W_{11}r_{11} + W_{12}r_{12} + W_{13}r_{13} + W_{14}r_{14}$$

where $$W_{10} = E_{10}/(E_{10} + E_{11} + E_{12} + E_{13} + E_{14}) \text{, etc.}$$

The estimates of R1_PR, R2_PR, and R3_PR, using the supplemented data, make the most complete possible use of the data by combining the women's survey with the household survey. They add exposure to the denominators of the rates, similar to how DHS supplements the exposure in a survey limited to ever-married women, as will be described in section 2.7. However, they require considerably more data manipulation than two alternatives that will be described in Sections 2.5 and 2.6. Moreover, as the most data-based approach, it will be unreliable if the data are of poor quality. The method depends on an assumption that age is reported correctly, both for girls/women who had an early birth and for those who did not.

2.5 Calculating Fertility Rates for Pooled Ages: "Lexis" Method

The procedure just described makes the most complete possible use of the available data for both the numerators and the denominators of the rates for age 12-14 and age 10-14. A simpler approach uses re-estimated weights derived from the Lexis diagram in Figure 2.1.

[9] We use the term "re-weight" rather than simply "weight" to distinguish from the use of the sampling weights, which are always included in the estimation.

Consider the pooled rate for age 12-14 during the 3 years before the survey. The box in Figure 2.1 that represents the relevant back-dated exposure of women who were age 15-19 at the time of the survey indicates that the data include 1/6 of the true or uncensored exposure to age 12 during the 3 years before the survey. That is, in the row for exact age 12, there are three small 1-year squares for the 3 years of exposure, and only half of one of the squares is shaded dark blue, amounting to 1/6 of the total area of the three small squares. Similarly, the exposure to age 13 is 3/6 of the total, and the exposure to age 14 is 5/6 of the total. These fractions are shown just to the right of the colored area in the Lexis diagram on the left side of Figure 2.1.

The approach uses factors *f* that will inflate the observed exposure *e* according to the geometry of the Lexis diagram and construct new weights that sum to 1. The inflation factors described in the preceding paragraph would be 6/1, 6/3, and 6/5, for ages 12, 13, and 14 respectively. We can ignore the 6 that is included in these multipliers because it would cancel out in the normalization of the new weights to sum to 1. Incorporating the factors 1, 1/3, and 1/5, the new weights would be

$$W_{12} = e_{12}/D$$

$$W_{13} = (e_{13}/3)/D$$

$$W_{14} = (e_{14}/5)/D$$

where the denominator D is $D = e_{12} + (e_{13}/3) + (e_{14}/5)$.

The pooled rate for age 12-14 in the past 3 years using these rates based on the Lexis diagram, that is, based on the assumption of a uniform distribution of exposure over time *within* each single year of age, and using the preceding weights, would be given by

$$R1\_Lexis = W_{12}r_{12} + W_{13}r_{13} + W_{14}r_{14}$$

The estimate of R2 (for age 10-14 in the past 3 years) is limited by the absence of any information about exposure (as well as births) for age 10-11 in the past 3 years. From the geometry of Figure 2.1, the most plausible assumption is that $P = \frac{3}{5} = 0.6$ of the exposure to age 10-14 was at age 12-14. Therefore,

$$R2\_Lexis = 0.6 * R1\_Lexis \ .$$

For the 5-year rate R3, the inflation factors for the five single-year rates will be proportional to 10/1, 10/3, 10/5, 10/7, and 10/9. Ignoring the factor of 10, the five weights would be given by

$$W_{10} = e_{10}/D$$

$$W_{11} = (e_{11}/3)/D$$

$$W_{12} = (e_{12}/5)/D$$

$$W_{13} = (e_{13}/7)/D$$

$$W_{14} = (e_{14}/9)/D$$

where the denominator *D* is $D = e_{10} + (e_{11}/3) + (e_{12}/5) + (e_{13}/7) + (e_{14}/9)$.

Using these weights, the estimated rate for age 10-14 in the past 5 years would be given by

$$R3\_Lexis = W_{10}r_{10} + W_{11}r_{11} + W_{12}r_{12} + W_{13}r_{13} + W_{14}r_{14} \ .$$

This estimator was described in Chapter 6 and Appendix 1 of DHS Comparative Report 45 (MacQuarrie, Mallick, and Allen 2017). The methods to calculate it are the same in this report as in Comparative Report 45.[10] It is based on an assumption that, within each of the 3 or 5 years of age that the respective rates refer to, the women who appear in the survey of women are representative of the girls/women who did not appear in the survey of women because they were too young at the time of the survey. Variation in exposure, from one year of age to another, is determined solely by the variations in the women's survey.

2.6 Calculating Fertility Rates for Pooled Ages: "Equal" Method

The pooling procedure described in Section 2.5 is based entirely on data obtained in the survey of women, but it allows the amount of exposure to vary across single years of age. An even simpler approach, referred to as the Equal method, does not use the Lexis diagrams in Figure 2.1 at all. It simply gives equal weight to each of the single-year rates, in effect approximating the distribution of exposure within age 12-14 or age 10-14 with a uniform distribution. The amount of exposure is assumed to be the same in each single year of age. In Section 2.5, a weak version of this assumption was used in the approximation $P = \frac{3}{5} = 0.6$.

With the Equal method, the formulas for R1_equal, R2_equal, and R3_equal are the same as given in section 2.5, but in the formula for R1_equal, each W is 1/3 and in the formula for R3, each W is 1/5. This method does not have a demographic rationale, but is included because it is simpler than the other two methods and we will find that it gives results that are surprisingly similar to those methods.

2.7 Modifications for Surveys of Ever-Married Women

As mentioned, some DHS surveys are limited to ever-married women and are referred to as EMW surveys. In such surveys, the calculation of fertility rates that can be interpreted as estimates for all women, regardless of marital status, require inflation of the exposure in the denominators. All of the files for women include "all-women-factors" with a variable "awfactt".[11] This factor is inversely proportional to the fraction of all women in the household roster (in each combination of sample stratum and single year of age) who are ever-married. In a non-EMW survey, awfactt=100 for all cases. In EMW surveys, awfactt/100 is the inflation factor for the contributions to the denominators. The factor can be large for women age 15, 16, etc. It declines steadily toward 1 as age increases and an increasing proportion of women are ever-married.

The standard approach to calculating fertility rates incorporates awfactt when it is needed. Exactly the same approach is extended to the fertility that a woman age 15 or older had before age 15. For the method based on the Lexis diagram, the exposure *e* for each age would be the exposure after inflation with the all-women factors. No other modifications are required.

[10] The authors of the present methodological report prepared Appendix 1 and the relevant under-15 fertility estimates that were included in Comparative Report 45. There is complete consistency between this report and Comparative Report 45 and the 2018 version of the Guide to DHS Statistics.
[11] The final "t" in "awfactt" indicates a version of "awfact" that applies to the total sample. Different final letters are used for factors that apply for covariates. For example, if fertility rates by wealth quintile are needed, the appropriate version would be "awfactw." This report refers only to "awfactt."

2.8 Standard Errors and Confidence Intervals

This report includes some use of 95% confidence intervals for the rates, adjusted for weights, clustering, and sample stratification, but we will not go into detail on how those confidence intervals are calculated. DHS provides confidence intervals for the total fertility rate, in Appendix B on sampling error in the main survey reports. That confidence interval is estimated with a jackknife approach. All estimates of demographic rates that appear in DHS main survey reports and on STATcompiler are calculated by dividing total births by total exposure, without a statistical model of any kind, using CSPro, and all the standard errors and confidence intervals for rates given in Appendix B are obtained in a separate package using the jackknife method.

In this report, we do not actually calculate the single-year rates by dividing a weighted number of births by a weighted number of woman-years of exposure. Rather, the single-year rates are calculated with a stacked Poisson regression, using individual women as cases, in which the number of births that the woman had in an interval of age and time is the outcome and the log of the woman's exposure is an offset. Adjustments for weights, clustering, and sample stratification are included in the model.[12] The coefficients produced by the model, including the variance-covariance matrix, are used to produce the single-year rates, their standard errors, and 95% confidence intervals for the rates. All the pooled estimates R1, R2, and R3 are weighted combinations of the single-year rates and all of them fit into a single statistical framework, the only differences being in the choice of weights.

Regardless of how the confidence intervals are obtained, they are wide for the under-15 fertility rates compared to the rates for later ages, in the sense that if L and U are the lower and upper ends of the confidence interval, respectively, the ratio U/L can be large. This can happen even when the arithmetic difference U-L is small, because the point estimates of the rates are so close to 0.

Confidence intervals for numbers close to 0, with a natural 0, such as the under-15 fertility rates, should be calculated in such a way that the lower end of the interval cannot be negative. The intervals used here are symmetric on a log scale. That is, if R is the point estimate of a rate, then log(R) is halfway between log(L) and log(U). In terms of ratios, U/R=R/L.

To give a specific example using the PR method, the estimated fertility rate R2 (for age 10-14 in the past 5 years) in the India 2015-16 survey is 0 births per 1,000 years of exposure. The rate is 0 after rounding, but there are indeed a few births before age 15, and the estimated rate with two decimal places is 0.18. The high and low values of a 95% confidence interval are 0.67 and 0.05, respectively. The width of the interval is only 0.67 - 0.05 = 0.62. However, the ratio of the high end to the low end is 0.67/0.05=14.26 (these calculations were made with more decimal places). When analyzing rare events, such as under-15 fertility, statistical inference must be undertaken with caution.

The standard error of a rate, or a log of a rate, is affected by the number of cases in the numerator—the observed number of births—not just the number of cases in the denominator—the number of women or women-years of exposure. Again, say that b is the number of births, e is the woman-years of exposure, and $r = b/e$ is the rate. Under the Poisson approximation, before any adjustments for weights, clustering, or stratification, the standard error of r will be directly proportional to r itself and inversely proportional to the

[12] All calculations were done in Stata, using data files publicly available at www.dhsprogram.com.

square root of b, because $se(r) = \frac{\sqrt{b}}{e} = r/\sqrt{b}$. (For a pooling of ages, this is an approximation that ignores variation across age.)

Using the Lexis diagrams in Figure 2.1, we can easily simulate the number of additional births that would be brought into the sample if, say, the lower age of eligibility were reduced. This simulation is useful partly for gauging the effect on the confidence interval, but also for describing the potential return for hypothetically revising the lower age boundary.

We simply apply the inflation factors given earlier to the observed numbers of births rather than to the woman-years of exposure. Define b_a to be the observed number of births in the past 5 years at age a, for a=10, 11, 12, 13, or 14, among the women who are age 15-19 at the time of the survey. Let B be the total number of births, i.e., $B = b_{10} + b_{11} + b_{12} + b_{13} + b_{14}$. If we had full exposure to those ages, the expected or simulated number of births would be

$$\hat{B} = (\tfrac{10}{1})b_{10} + (\tfrac{10}{3})b_{11} + (\tfrac{10}{5})b_{12} + (\tfrac{10}{7})b_{13} + (\tfrac{10}{9})b_{14}, \text{ or}$$

$$\hat{B} = 10(b_{10} + b_{11}/3 + b_{12}/5 + b_{13}/7 + b_{14}/9).$$

The ratio $\hat{B}/B$ would be the multiplier to the number of births gained by dropping the minimum age to 10. Although it is always greater than 1, it is typically much closer to 1 than might be expected. The percentage reduction in the width of the confidence interval for R3 would be approximately $100 * [1 - sqrt\left(\frac{B}{\hat{B}}\right)]$. This report will provide estimates of the additional number of births and the multiplier if the minimum age were reduced to 14 or 13.

2.9 Example: The Mali 2012-13 Survey

We will illustrate the steps for estimating the pooled under-15 fertility rates with a specific survey, the 2012-13 DHS survey of Mali. This survey is selected because it has the highest under-15 fertility in the past 5 years among the 67 surveys that will be analyzed in Chapter 3.

The rates for the 3 years before the survey will be described first, using Table 2.1, and then the rates for the 5 years before the survey, using Table 2.2.

The column headings for these tables are described in the following list:

Column 1: Age in single years

Column 2: Births at the specified age, from the women's data

Column 3: Years of exposure, from the women's data

Column 4: Revised years of exposure, women's data supplemented with household data

Column 5: The single-year fertility rate (1000*Column 2 / Column 3); bottom entry is the raw rate

Column 6: Weights (proportions) from supplemented data

Column 7: Column 3, times the factors 1, 1/3, and 1/5 (Lexis diagram weights before normalization)

Column 8: Weights (proportions) from Lexis diagram

Column 9: Weights (proportions) from assumption of uniform distribution

Column 10: Column 4 * Column 6; bottom entry is R1_PR from Section 2.4

Column 11: Column 4 * Column 7; bottom entry is R1_Lexis from Section 2.5

Column 12: Column 4 * Column 8; bottom entry is R1_equal from Section 2.6

Table 2.1 Worksheet to calculate fertility rates for single years of age 12, 13, and 14 and pooled rates for 12-14 in the past 3 years with alternative methods. Mali 2012-13 DHS

Col. 1	Col. 2	Col. 3	Col. 4	Col. 5	Col. 6	Col. 7	Col. 8	Col. 9	Col. 10	Col. 11	Col. 12
Age	Births	Exposure	Expos.rev	Fert.rate	W_PR	Expos.*f	W_Lexis	W_Equal	R1_PR	R1_Lexis	R1_equal
10	0.00	0.00	2429.18		0.00	0.00	0.00	0.00			
11	0.00	0.00	2406.82		0.00	0.00	0.00	0.00			
12	0.00	159.24	2226.57	0.00	0.42	159.24	0.30	0.33	0.00	0.00	0.00
13	9.93	553.38	1757.81	17.94	0.33	184.46	0.35	0.33	5.98	6.32	5.98
14	33.32	898.41	1292.46	37.09	0.24	179.68	0.34	0.33	9.08	12.73	12.36
	43.25	1611.04	10112.85	26.85	1.00	523.39	1.00	1.00	15.06	19.06	18.34

The data underlying Table 2.1 are provided in Columns 2, 3, and 5. There are no births for ages 10 and 11 because there was no exposure to these ages, as described earlier. It happens that this survey also had no births at age 12 during the previous 3 years, but that is an empirical zero rather than a structural zero.

In an EMW survey, the total person-years of exposure given in Column 3 would be the expanded number of woman-years, following application of awfactt/100.

The weighted total number of births during age 12-14 was only 43.25, with 1,611.04 woman-years of exposure. The ratio of these two numbers, multiplied by 1,000, is the pooled raw rate, 26.85 births per 1,000 women or per 1,000 woman-years of exposure to age 12-14. The age-specific rates and the raw rate are given in Column 5. The methods in this chapter are intended to improve upon the raw rate at the bottom of Column 5 by taking into account the censoring due to the age criterion for eligibility.

The supplemented woman-years of exposure in Column 4 consist of three pieces: the exposure from the women's data in Column 3; the exposure from girls age 10-14 in the household survey who were living with their mother, and therefore appeared in the birth histories and had a cmc of birth; and the exposure from other girls age 10-14 in the household survey whose cmc of birth was imputed by setting their most recent birthday at six months before the month of interview. Column 5 includes entries for ages 10 and 11 as well as age 12-14. The total for age 12-14 was 5,276.84 woman-years; the total for age 10-14 was 10,112.85; and the ratio of the former to the latter is the proportion P=0.5218. The calculation of this proportion is the only use of any data for age 10-11 in the past 3 years. P will be used to expand the rate for age 12-14 to a rate for 10-14.

Column 6 gives the proportion of the supplemented exposure for age 12-14 that is within each single year of age; Column 8 gives the corresponding weights or proportions for the Lexis diagram approach; and

Column 9 gives the proportions under the assumption of a uniform age distribution. Columns 10, 11, and 12 give the products of Column 4 and Columns 6, 8, and 9, respectively. The totals for the last three columns are the estimates R1 under the three approaches.

In the example, Column 6 shows that the proportions of exposure to age 12-14 that is in the single years 12, 13, and 14 are 42%, 33%, and 24% respectively. The corresponding allocation from the Lexis diagram and the denominators of the age-specific rates, given in Column 7, are 30%, 35%, and 34% respectively. The uniform allocation given in Column 9 is 33%, 33%, and 33% respectively.

Column 7 is included in the worksheet because it provides an intermediate step to obtaining Column 8. It is calculated as the product of the observed exposure in Column 3 and the age-specific factors from the Lexis diagram (1 for age 12, 1/3 for age 13, and 1/5 for age 14).

Column 8 is proportional to Column 7 and is obtained by normalizing the entries in Column 7 to add to 1 (that is, dividing the numbers in Column 7 by the total for Column 7).

The estimates of the pooled fertility rate for age 12-14 in the past 3 years are as follows:

R1_PR from Column 10, which makes maximum use of all the data, is 15.06.

R1_Lexis from Column 11, which uses weights implied by the Lexis diagram, is 19.06.

R1_equal from Column 12, a simple average of the three single-year rates, is 18.34.

R1_PR is the lowest of these three estimates because the exposure distribution in Column 6 is weighted toward the single-year rate for age 12, which is the lowest of the single-year rates. R1_Lexis and R1_PR are similar to each other because the exposure distributions in Column 8 are close to the uniform distribution in Column 9. Later, we will comment more on these three distributions of exposure.

It is possible to produce estimates of R2, the pooled rate for age 10-14 in the past 3 years, from R1, by assuming that fertility at age 10-11 is negligible, as follows (the estimates of R2 are not explicitly included in Table 2.1):

R2 from Column 10 uses a multiplier of P=0.5218: R2_PR=P*R1_PR=7.86.

R2 from Column 11 uses a multiplier of P=0.6: R2_Lexis=P*R1_Lexis=11.43.

R2 from Column 12 also uses a multiplier of P=0.6: R2_equal=P*R1_equal=11.01.

We next turn to the calculation of the three estimates of R3. Table 2.2, which contains the necessary summary data for R3, is similar to Table 2.1 and requires little additional discussion. This table refers to the 5 years before the survey and age 10-14. The absence of any births at age 10-11 is an empirical finding, not an artifact of the data structure.

Again, Column 7 provides an intermediate step for obtaining Column 8. It is calculated as the product of the observed exposure in Column 3 and the age-specific factors from the Lexis diagram on the right side of Figure 2.1 (1 for age 10, 1/3 for age 11, 1/5 for age 12, 1/7 for age 13, and 1/9 for age 14). Column 8 is proportional to Column 7 and is obtained by normalizing the entries in Column 7 to add to 1 (that is, dividing the numbers in Column 7 by the total for Column 7).

Table 2.2 Worksheet to calculate fertility rates for single years of age 10, 11, 12, 13, and 14 and pooled rates for 10-14 in the past 5 years with alternative methods. Mali 2012-13 DHS

Col. 1	Col. 2	Col. 3	Col. 4	Col. 5	Col. 6	Col. 7	Col. 8	Col. 9	Col. 10	Col. 11	Col. 12
Age	Births	Exposure	Expos.rev	Fert.rate	W_PR	Expos.*f	W_Lexis	W_Equal	R3_PR	R3_Lexis	R3_equal
10	0.00	159.24	3792.86	0.00	0.26	159.24	0.18	0.20	0.00	0.00	0.00
11	0.00	553.38	3354.26	0.00	0.23	184.46	0.20	0.20	0.00	0.00	0.00
12	5.30	898.41	2965.73	5.90	0.20	179.68	0.20	0.20	1.18	1.17	1.18
13	22.82	1322.41	2526.84	17.26	0.17	188.92	0.21	0.20	2.95	3.60	3.45
14	100.35	1730.33	2124.38	57.99	0.14	192.26	0.21	0.20	8.34	12.33	11.60
	128.47	4663.78	14764.07	27.55	1.00	904.56	1.00	1.00	12.48	17.10	16.23

The relevant estimates of the rates for age 10-14 in the past 5 years are given in the bottom row for Columns 5, 10, 11, and 12. The estimate from Column 5, with no adjustments at all, is 27.55. It is obvious that this rate is much too high, because it includes so little observed exposure to the younger ages 10-11 in Column 3.

R3_PR from Column 10, which makes maximum use of all the data, is 12.48.

R3_Lexis from Column 11, which uses weights implied by the Lexis diagram, is 17.10.

R3_equal from Column 12, a simple average of the five single-year rates, is 16.23.

The purpose of this example has been to demonstrate the data and the steps that go into the calculations, but there are some potentially generalizable inferences as well. The evidence from this example is that all three estimates of R1 are close to one another; all three estimates of R2 are close to one another; and all three estimates of R3 are close to one another—in terms of the number of births per 1,000 women or woman-years of exposure. The Lexis and Equal estimates are closest. In relative terms, in this example, the PR estimates are about 20% to 30% lower than the other two estimates.

The assessment that two estimates are "close" can be based on their arithmetic difference or their relative difference. It must be kept in mind that all of these rates are very small compared with the rates for the later age intervals. In the Mali 2012-13 survey, the estimated 3-year fertility rate on STATcompiler for age 15-19 is 172, and for age 25-29 (the peak age interval) it is 272 per 1,000 woman-years. When compared with rates of that magnitude, alternative rates for age 10-14 that differ by, say, five births per 1,000 woman-years can be described as "close". When comparing small numbers, the arithmetic difference is more descriptive than the relative difference.

The raw or simple estimate is always too high because it is biased toward age 14. In this example, the PR estimate is always too low, because it uses an age distribution of single years of age within the range 10-14 that is implausibly tilted toward the early ages, as will be discussed further in Chapter 3. The Lexis and Equal estimates are always between the raw and PR estimates and close to each other.

Another useful comparison is between R2 and R3. These are the pooled estimates for ages 10-14 based on the past 3 years and the past 5 years, respectively. The estimates from the supplemented data are 7.86 and 12.48, respectively. The estimates using the Lexis diagram are 11.43 and 17.10, respectively. The estimates using the Equal assumption are 11.01 and 16.23, respectively. Because the time intervals are different, we would not expect exact agreement—but because the data for the past 3 years are the majority of the data for the past 5 years, we would expect general agreement. The alternative estimates are not as close as we might expect, at least for this example. The difference is about five births per 1,000; all the R2 estimates are about one-third below the corresponding R3 estimates.

On STATcompiler, the values of R2 and R3 are given as 11 and 17, respectively. These are the rounded values of R2_Lexis=11.43 and R3_Lexis=17.10.

If the age of eligibility for this survey had been age 10, rather than 15, then the (weighted) number of births at age 10-14 in the past 5 years would have been more than the 128.47 given in the bottom row of Column 2 in Table 2.2. However, most of the increase in exposure would have been to the youngest ages, for which fertility is low, so the gain in the number of births would be small. Using the formula given in Section 2.8, we estimate that the observed number of births would have increased by only about 20%, to 154.69. The reduction to the width of a confidence interval for the R3 estimates would have been approximately 9%, surprisingly little gain for adding a large age group to the survey of women.

We now turn to calculations of these estimates with a large number of surveys and an examination of their correspondences. As will be seen later, the Mali survey has the largest discrepancies between alternative estimates of all the surveys examined in this report.

3 ESTIMATES OF UNDER-15 FERTILITY RATES

The previous chapter described how censoring affects estimates of the single-year under-15 fertility rates and the pooled rates R1, R2, and R3. The effect on the single-year rates was judged to be negligible and any procedures to adjust for it would fall outside the scope of this report. Three procedures were proposed to correct for the effect of censoring on the pooled rates. The first procedure involves supplementing the women age 15-19 in the women's survey with girls age 10-14 in the household survey and then re-estimating exposure to ages 10-14. This approach makes the most complete use of the available data, but it requires more data processing and is not necessarily optimal if it differs negligibly from simpler alternatives. The second procedure is based on a visual analysis of the Lexis diagram in Figure 2.1. The second procedure is simpler than the first, and for that reason may be preferable if the numerical results of the two methods tend to be similar. The third procedure, with equal weighting of the single-year rates, is even simpler, but does not have a demographic rationale.

The three methods for pooled rates simply re-weight the rates for single years of age. The standard errors of all the rates depend ultimately on the numbers of births and we do not expect noticeable differences between the three types of pooled rates in their standard errors or confidence intervals, but that issue will be taken into account. Confidence intervals can be estimated for the pooled rates as well as the single-year rates.

3.1 Data

The empirical analysis is based on all countries with at least one standard DHS survey since 2000.[13] If there is more than one such survey, only the most recent one is included. The most recent available rounds of the Continuous Surveys in Peru and Senegal are included. Malaria Indicator Surveys (MIS), AIDS Indicator Surveys (AIS), and Special DHS surveys are excluded, however, because they generally have no birth histories or have truncated birth histories. There are also some differences in methodology and interviewer training with respect to the birth histories, in particular, that cast doubt on comparability. For example, in MIS surveys the birth history generally begins with the most recent birth, rather than the first birth, besides being truncated. These criteria led to the selection of 67 surveys in 67 countries.[14] Appendix Table A1 lists all the surveys in our analysis, including some that recorded no births at all before age 15. That table includes the single-year rates.

Because most of the countries that conducted a DHS survey since 2000 had several such surveys, the earliest date on the list is 2001, and only 13 of the most recent surveys were conducted before 2010. The median survey year is 2013, and the most recent survey year is 2016.

This list includes seven EMW surveys (Afghanistan 2015, Bangladesh 2014, Egypt 2014, Jordan 2012, Maldives 2009, Pakistan 2012-13, and Turkey 2003). One survey, Colombia 2015, included girls age 13 and 14 in the survey of women. The analysis will impose a strict lower age boundary of 15. Girls age 13 and 14 in the Colombia survey will be dropped from the survey of women but included in the household

[13] In order to be included, the data files had to be available on the DHS website by July 31, 2018.
[14] The most recent survey of Vietnam was conducted in 2002. That survey met the criteria but was omitted because it did not include a code (b16) that is required to link the birth histories with the household file.

survey. Section 3.5 will include some discussion that is specific to this survey, and the gain from including age 13 and 14 in the survey of women.

3.2 Summary of Estimates

Appendix Table A2, which like Appendix Table A1 provides a list of all the surveys in this report, includes the estimates developed in Chapter 2, each of which is a weighted pooling of the single-year age-specific rates given in Appendix Table A1:

R1: the fertility rate for girls age 10-12 in the past 3 years

R2: the fertility rate for girls age 10-14 in the past 3 years

R3: the fertility rate for girls age 10-14 in the past 5 years

For each rate we provide three alternatives:

PR: the estimate that includes full exposure from the PR file of household members

Lexis: the estimate that does not use the PR file, but simulates full exposure from a Lexis diagram

Equal: the estimate that simply averages the age-specific rates for the past 3 or 5 years

We also calculate P: the percent of exposure to age 10-14 in the past 3 years that was to age 12-14.

Chapter 2 included details on the nine possible calculations of weighted pooling and an example worked out from the Mali 2012-13 DHS survey. All the calculations include many decimal places and the tables in Appendix A include extra decimal places simply to enable a reader to replicate the calculations. Table A1 gives the single-year age-specific rates that underlie the pooled rates, and the pooled rates are given in Table A2.

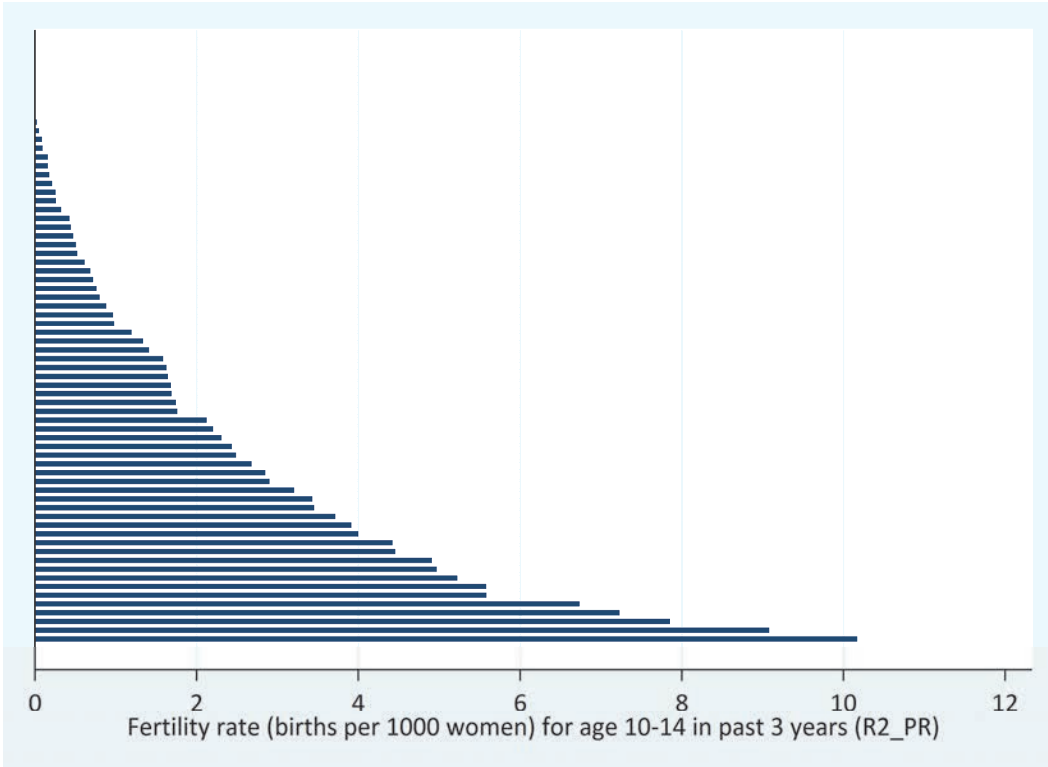
The rates are expressed as the number of births per 1,000 women (or, more accurately, per 1,000 woman-years of exposure). The detail provided in Appendix A and in the example goes far beyond a useful level of precision. Most of the rates are small by any standard.

Table 3.1 lists the surveys for which any of the nine rates rounded to a whole number greater than 0 births per 1,000. The table includes 53 surveys, ranked by the magnitude of R2_PR. A graphical overview is provided with Figure 3.1, a horizontal bar graph that gives all 67 values of R2_PR. Only a handful of the 67 countries have R2_PR values above 6 per 1,000; the maximum value is 11. The highest countries/surveys are Madagascar 2008, Guinea 2012, Mali 2012, Angola 2015, and Bangladesh 2014.

Table 3.1 Alternative estimates of fertility rates (births per 1,000 years of exposure) for age 12-14 in the past 3 years (R1), age 10-14 in the past 3 years (R2), and age 10-14 in the past 5 years (R3), ranked by R2_PR and rounded to the nearest integer.

	Year	Estimates of R1			Estimates of R2			Estimates of R3		
Country		PR	Lexis	Equal	PR	Lexis	Equal	PR	Lexis	Equal
Afghanistan	2015	0	0	0	0	0	0	1	1	1
Turkey	2003	0	0	0	0	0	0	1	1	1
India	2015	0	1	1	0	1	1	0	0	0
Timor-Leste	2016	1	1	1	0	0	0	0	0	0
Egypt	2014	1	1	1	0	0	0	1	1	1
Burkina Faso	2010	1	1	1	0	1	1	1	1	1
Nepal	2016	1	1	1	0	1	1	1	1	1
Ethiopia	2016	1	1	1	1	1	1	1	1	1
Tanzania	2015	1	1	1	1	1	1	1	1	1
Burundi	2016	1	1	1	1	1	1	1	1	1
Myanmar	2015	1	1	1	1	1	1	1	1	1
Rwanda	2014	1	1	1	1	1	1	0	0	0
Philippines	2013	1	1	1	1	1	1	1	1	1
Lesotho	2014	1	2	2	1	1	1	1	1	1
Guyana	2009	2	2	2	1	1	1	3	3	3
Yemen	2013	2	2	2	1	1	1	1	1	1
Ghana	2014	2	2	2	1	1	1	1	1	1
São Tomé and Principe	2008	2	3	3	1	2	2	1	2	2
Haiti	2016	2	2	2	1	1	1	1	1	1
Senegal	2016	2	3	3	1	2	2	2	2	2
Zambia	2013	3	3	4	2	2	2	2	2	2
Gambia	2013	3	3	3	2	2	2	2	3	3
Zimbabwe	2015	3	3	3	2	2	2	1	1	1
Kenya	2014	3	3	3	2	2	2	2	2	2
Comoros	2012	3	3	3	2	2	2	2	2	2
Uganda	2016	3	3	3	2	2	2	2	2	2
Togo	2013	3	3	4	2	2	2	2	3	3
Benin	2011	4	4	4	2	2	2	3	3	3
Dominican Republic	2013	4	4	4	2	2	2	4	4	4
Bolivia	2008	4	5	5	2	3	3	2	3	3
Colombia	2015	4	4	4	2	2	3	3	3	3
Malawi	2015	5	5	5	2	3	3	2	3	3
Liberia	2013	5	5	5	3	3	3	5	5	5
Congo Democratic Republic	2013	5	6	6	3	4	4	5	6	5
Nigeria	2013	5	5	6	3	3	3	4	4	5
Swaziland	2006	6	6	6	3	3	4	2	3	3
Guatemala	2014	6	6	6	3	3	4	3	3	3
Honduras	2011	6	6	6	3	4	4	4	4	4
Niger	2012	8	9	10	4	6	6	7	9	9
Namibia	2013	7	7	7	4	4	4	2	2	2
Nicaragua	2001	7	7	7	4	4	4	4	4	4
Congo	2011	8	8	8	4	5	5	5	5	5
Côte d'Ivoire	2011	8	9	9	4	5	6	6	7	7
Sierra Leone	2013	8	8	9	5	5	5	6	6	7
Chad	2014	9	10	11	5	6	6	9	10	11
Mozambique	2011	10	9	10	5	6	6	6	6	6
Cameroon	2011	10	10	11	6	6	6	7	8	8
Gabon	2012	10	11	11	6	7	6	6	6	6
Madagascar	2008	12	13	13	7	8	8	8	9	9
Guinea	2012	14	14	15	7	9	9	9	10	11
Mali	2012	15	19	18	8	11	11	12	17	16
Angola	2015	16	18	18	9	11	11	9	11	11
Bangladesh	2014	17	17	17	10	10	10	9	9	9

Figure 3.1 Bar graph of the distribution of the fertility rate (births per 1,000 woman-years of exposure) for age 10-14 in the past 3 years (R2_PR). Includes all 67 surveys.



3.3 Comparisons of Estimates

This section compares the estimated fertility rates using PR weights, Lexis weights, and equal weights, respectively. Even a cursory examination of Table 3.1 shows a high level of correspondence among the three versions of R1, R2, and R3. The most important comparisons are for R2 and R3, because they are the most direct extension of the standard rates given in all of the main reports. We are also most interested in comparisons with the Lexis estimates because they have already appeared in Comparative Report 45 and are in place on STATcompiler.

The detailed comparisons will be based on three criteria: nearness in terms of the absolute arithmetic difference; nearness in terms of the absolute relative difference; and whether each estimate is within a 95% confidence interval for the other estimate. The comparisons will be done before rounding, although some of the estimates given in tables and the text will be rounded to the nearest integer. The first two criteria are applied in Figures 3.2, 3.3, and 3.4.

Say that two rates being compared are Ra and Rb. We prefer comparisons that are symmetric with respect to which rate is labelled Ra and which is labelled Rb. That is, if D is the measure of the difference between Ra and Rb, we prefer that D(Ra,Rb)=D(Rb,Ra). The absolute arithmetic difference, D=abs(Ra-Rb), has this property. It is expressed as a difference in births per 1,000 women, ignoring the direction of the difference. To describe the relative difference, we use a log scale and define the absolute relative difference to be D=100*abs[log(Ra/Rb)]. This can be interpreted as the percentage relative difference between the two estimates, on a log scale, ignoring the direction of the difference.

Figure 3.2 contains six subfigures describing the estimates of R1, the fertility rate for age 12-14 in the past 3 years, arranged in three rows and two columns. The first row compares R1_Lexis with R1_PR; the second row compares R1_equal with R1_PR; the third row compares R1_equal with R1_Lexis. The scatterplots in the left column show the surveys as points and a straight line of equality. The horizontal and vertical axes are the levels of the pairs of estimates. Our interest is in how far the points are from the line of equality.

In the scatterplots on the right, the surveys are points, the horizontal axis is the measure of difference, and the vertical axis is the measure of relative difference. A vertical red line at an absolute arithmetic difference of 1 point marks a tolerance threshold of 1 birth per 1,000 in the estimate. Any dots to the right of the vertical line are flagged in terms of the arithmetic difference. A horizontal red line at an absolute relative difference of 10 marks a tolerance threshold of 10%. Any dots above the horizontal line are flagged in terms of the relative difference. Our criterion for flagging a pair of estimates on both criteria is that the dot for the survey is in the upper right quadrant of the scatterplot on the right. The most important comparisons are with the Lexis estimate, i.e., the comparisons in the first row and the third row of the figure.

Figures 3.3 and 3.4 are similar in structure but refer to R2 (the fertility rate for age 10-14 in the past 3 years) and R3 (the fertility rate for age 10-14 in the past 5 years), respectively. All three figures have the same scales on the vertical and horizontal axes for the corresponding subfigures. The India survey is excluded from the subfigures in the second column of Figures 3.2, 3.3, and 3.4 because it is an outlier with very high relative differences even though all the rates are extremely small.

Figure 3.2 Scatterplots of alternative estimates of R1, the fertility rate for age 12-14 in the past 3 years, showing absolute arithmetic differences and absolute relative differences

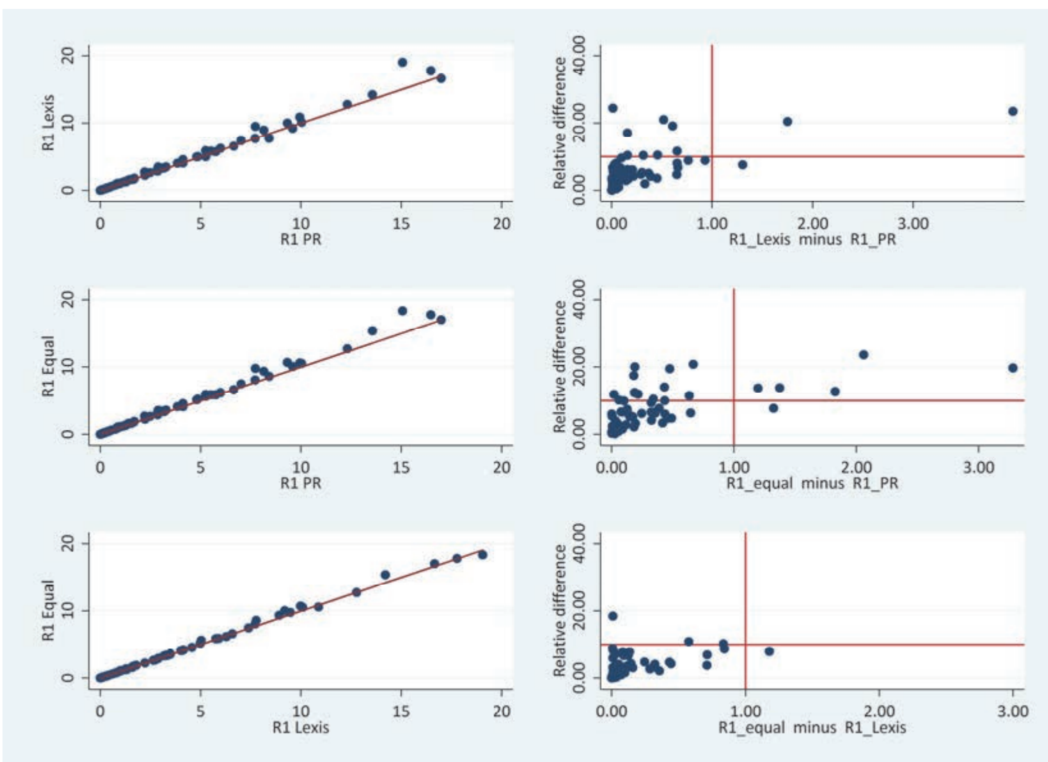


Figure 3.3 Scatterplots of alternative estimates of R2, the fertility rate for age 10-14 in the past 3 years, showing absolute arithmetic differences and absolute relative differences

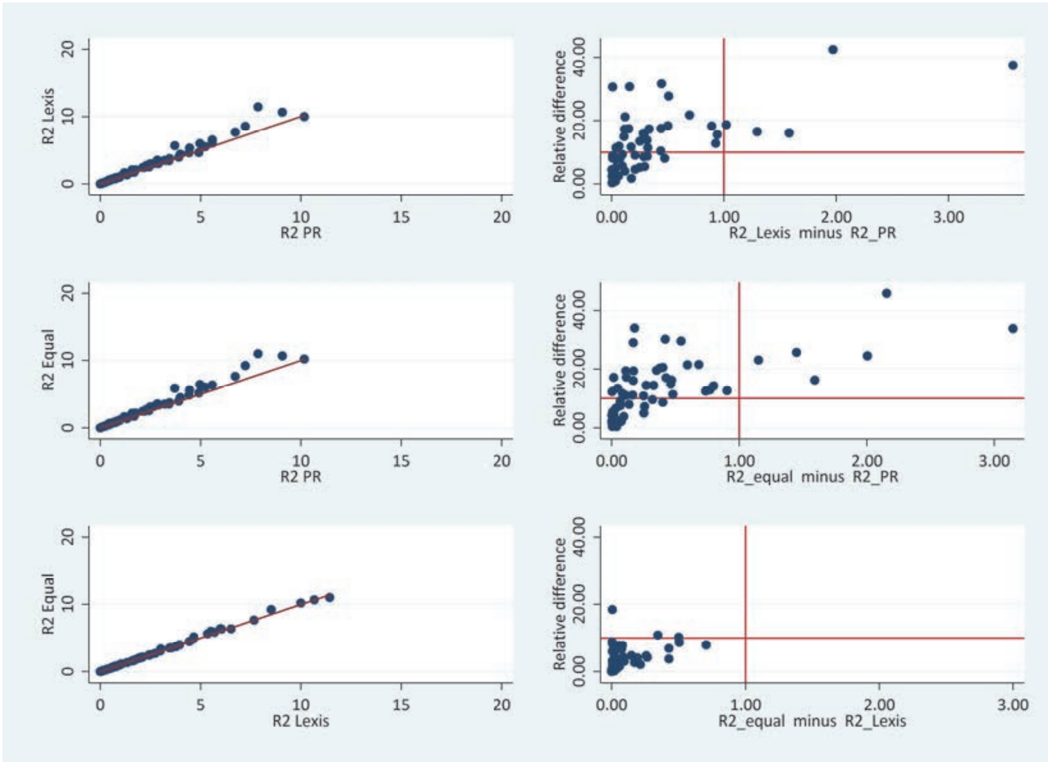
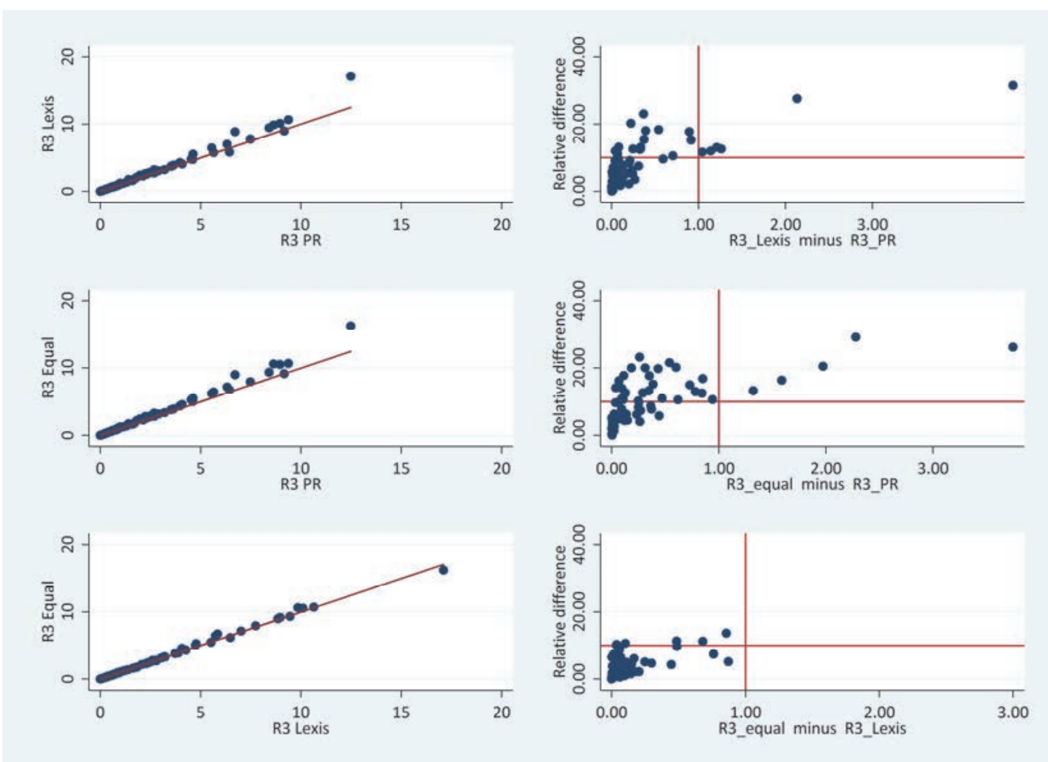


Figure 3.4 Scatterplots of alternative estimates of R3, the fertility rate for age 10-14 in the past 5 years, showing absolute arithmetic differences and absolute relative differences



Thresholds of 1 for an arithmetic difference and 10% for a relative difference are very stringent for rates that are so small in magnitude. We only flag the surveys that are outside both of these tolerances. In Figures 3.2, 3.3, and 3.4 these are the surveys that correspond with the dots in the upper right sector or quadrant of the subfigures in the second columns. There are only seven such surveys, and they appear repeatedly in those two sectors. Madagascar appears once, Côte d'Ivoire appears twice, Guinea four times, Chad and Niger five times each, and Angola and Mali six times each. In magnitude, the most extreme differences are always for Mali. Angola and Mali have the highest levels of the under-15 rates, as well as the highest discrepancies among the alternatives. Four of the seven countries—Chad, Guinea, Mali, and Niger—were classified in Methodological Report 19 (Pullum and Staveteig 2017) in the poorest quintile of age and date quality in almost exactly the same set of 67 countries, using all surveys conducted from 2000 to 2015.

Unexpectedly perhaps, a simple averaging or equal weighting of the age-specific rates, as described in the bottom row of subfigures in Figures 3.2, 3.3, and 3.4, is very close to the estimates obtained from the Lexis diagram. Only one survey (Guinea, in the bottom row of Figure 3.2, for R1) is flagged in terms of the arithmetic difference and it is not flagged in terms of the relative difference, so there is not a single survey in the upper right quadrant for that comparison. All of the flagged combinations involve a comparison that involves the PR estimate.

The third comparison between estimates is in terms of whether the alternative point estimates are within one another's 95% confidence intervals. This is in no sense a statistical test, because the only difference between alternative (pooled) estimates is in the choice of weights *W*. However, a criterion for consistency can reasonably be based on the sampling variability of the estimates. It happens that such a criterion is always satisfied. Consider, for example, the discrepancy between R2_Lexis and R2_PR, which we have framed as one of the most important comparisons in this report. The survey with the largest discrepancy is Mali 2012-13, for which R2_PR=7.86 and R2_Lexis=11.43. The 95% confidence interval for R2_PR extends from 5.32 to 11.62 (as described earlier, it is symmetric on a log scale). The interval includes R2_Lexis=11.43. Similarly, a 95% confidence interval for R2_Lexis, which extends from 7.83 to 16.68, includes R2_PR=7.86. In both cases, the point estimate is close to an end of the other confidence interval, but is within the interval.

To summarize the comparisons, seven surveys are flagged in terms of differences of more than 1 birth per 1,000 or 10%, the first two criteria. In alphabetical order, Angola, Chad, Guinea, Mali, and Niger account for 27 out of 30 problematic comparisons. Niger has the largest discrepancies. After we apply the third criterion, expressed with confidence intervals, all of the comparisons could be classified as acceptable. The Lexis and Equal methods produce almost exactly the same estimates.

We will briefly comment on the likely reasons for the discrepancies that are detectable with the first and second criteria, even if not the third, continuing with the Mali survey as the main example. We attribute the larger discrepancies for Mali to a combination of three factors. The first, mentioned before, is that this survey shows relatively high under-15 fertility. Measurement errors usually tend to scale up in proportion to the level.

The second factor is that this survey shows a relatively large difference between the observed value of P and 0.6, the proportion of exposure to age 10-14 that is concentrated in age 12-14. The proportion calculated from the supplemented data is 0.5218, which is considerably below the value of 0.6 assumed with the Lexis and Equal methods. Thus R1_PR is about 20% below the other two estimates of R1, but the R2_PR is about 30% below the other two estimates of R2. That is, the smaller multiplier for the conversion from R1 to R2, for the PR-based estimate, exacerbates the discrepancy observed for the estimates of R1. The next section will explore this effect further.

The most important reason for the differences between alternative estimates in Mali is the differences in the age distributions of exposure shown in Columns 6, 8, and 9 of Tables 2.1 and 2.2. For the PR method, the age distributions are excessively skewed. In Table 2.1, the percentages of exposure at ages 12, 13, and 14 are 42%, 33%, and 24%, respectively. In Table 2.2, the percentages at ages 10, 11, 12, 13, and 14 are 26%, 23%, 20%, 17%, and 14%. Both of these distributions are probably weighted too heavily toward the younger ages, for which fertility is low.

By contrast, the Lexis-based age distribution of exposure, shown in Column 8 of Tables 2.1 and 2.2, shows an increase in exposure with age. The gradient is not steep, but there is more emphasis on the later ages than on the earlier ages. The equal-based distribution, with the same percentage at each age, is close to the Lexis-based distribution, yielding estimates that are almost identical.

The true age distribution of exposure within ages 12-14 and 10-14 in the past 3 years or the past 5 years would almost certainly follow the general shape of the age distribution, with more exposure in the younger ages and less in the older ages, but in a country with high fertility and mortality such as Mali the relative decline from one year of age to the next would be about 4%. Under this demographic scenario, a plausible distribution of exposure to ages 12, 13, and 14 would be 35%, 33%, and 32%, respectively. A plausible distribution of exposure to ages 10, 11, 12, 13, and 14 would be 22%, 21%, 20%, 19%, and 18%, respectively.

In a more detailed analysis of the Mali data that will not be presented here, it appears that there are substantial irregularities in the reported ages, suggesting both digit preference and transfers across the age 15 boundary that translate into irregularities in the backdated exposure. Misstatement of age in the vicinity of 15 may also tend to be different for girls/women who had an early birth than for girls who did not. Both the steep downward gradient in the distribution of exposure with the PR method (as age increases) and the gradual upward gradient with the Lexis method are implausible. The equal weights of the third alternative may actually be closest to the true distribution of exposure.

3.4 The Effect of Variation in the Age Distribution within Age 10-14

All of the exposure to age 10-14 in the past 3 years comes from women age 15-17 at the time of the survey, and when they are backdated by 3 years their exposure is entirely to age 12-14. Therefore, all three versions of R2—the fertility rate for age 10-14 in the past 3 years—required an assumption that fertility at age 10-11 is negligible and can be ignored. This assumption is supported by the single-year results presented in Comparative Report 45. Even the number of births observed at age 12 is negligible.

The transition from the three versions of R1—the fertility rate for age 12-14 in the past 3 years—to the respective versions of R2 requires an expansion of the age interval to encompass age 10-11. As described earlier, R2_Lexis and R2_equal are obtained from the geometry of the Lexis diagram, in which 3/5 of a surviving cohort's exposure to age 10-14 is to the last 3 years (age 12-14) and 2/5 is to the first 2 years (age 10-11). Thus, if the rate for age 10-11 is 0, we have R2_Lexis = 0.6*R1_Lexis + 0.4*0 = 0.6*R1_Lexis.

Similarly, we assumed that R2_equal = 0.6*R1_equal (but for R2_equal, the use of a factor of 0.6 does not really require reference to the Lexis diagram).

Figure 3.5 Histogram of the observed value of P, the proportion of exposure to age 10-14 in the past 3 years that is exposure to age 12-14, in the 67 surveys

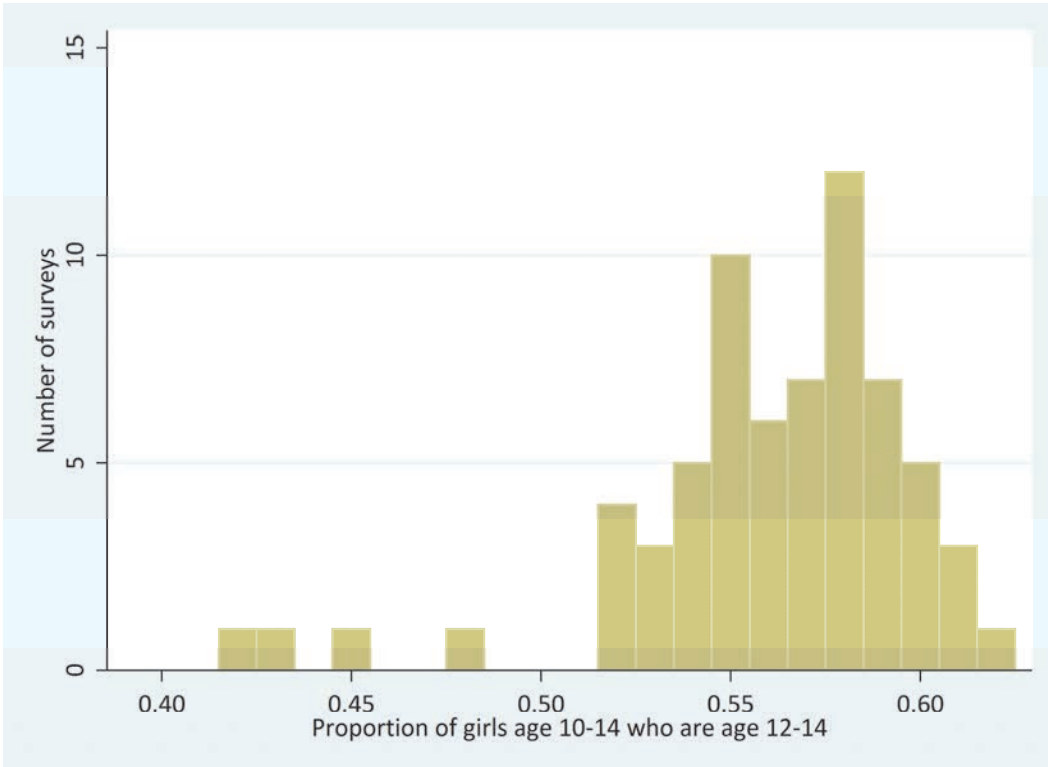


Figure 3.5 shows the distribution of the observed value of P in all the surveys. The lowest value is 0.4238 from the India survey and the largest is 0.6204 for Moldova. The overall mean is 0.5606. Rather than serving as an approximate mean of P, 0.6 is effectively an upper limit, because of the empirical tendency noted above for there to be a tilt toward the younger ages within the range of 12-14 or 10-14.

When P is actually less than 0.6, using 0.6 as the multiplier will give a value of R2_Lexis that is too high. Given other sources of error, empirical variation around an assumed value of P=0.6 is usually not serious enough to justify the effort required to refine the estimates of the fertility rates calculated with P=0.6, especially if under-15 fertility is very low, but for about a quarter of the surveys P is 0.54 or less, and the application of 0.6 as a default value leads to a non-trivial difference of about 10% or more. Of course, we do not know the true value and when the age data are of poor quality we may actually induce error by taking the data at face value and using the observed P in place of a default. Nevertheless, a default value of 0.6 is probably somewhat too high.

3.5 Simulation of a Reduction in the Lower Age of Eligibility

As described in Section 2.8, we can use the Lexis diagrams in Figure 2.1 to estimate the additional numbers of under-15 births in the past 3 years or 5 years that would have been identified in each survey if the minimum age for eligibility had been reduced below age 15. The option described in Section 2.8 was a hypothetical reduction to age 10. However, an average of fewer than two births was identified at age 10-12 in the past 5 years in all 67 surveys (the total weighted number in 67 surveys is 118.15 births, and all but a handful are at age 12). We will consider just two potential alternatives—a reduction to age 14 and a reduction to age 13, that is, a one-year reduction and a two-year reduction in the minimum age for eligibility. What would be the gain in terms of an improved estimate of the pooled fertility rates? Our approach to this question will assume that the point estimates of the single-year fertility rates and the pooled rates for age intervals would be unaffected. The question is simply about the degree to which the estimates would have higher precision in terms of the size of the confidence interval because of the increased number of births.

Consider first a reduction of the minimum age for eligibility to age 14 and the effect on the number of births in the past 3 years. Going back to the Lexis diagram on the left side of Figure 2.1, this change would have the effect of lowering, by one year, the diagonal line that separates the blue and red portions of the Lexis diagram. The blue area for age 14 would increase from 2.5 units to 3.0. The blue area for age 13 would increase from 1.5 units to 2.5. The blue area for age 12 would increase from 0.5 units to 1.5. The blue area for age 11 would increase from 0.0 units to 0.5, but we have no data for age 11 so it must be ignored. We can apply inflation factors to the numbers of births observed at ages 14, 13, and 12. Based on the increase in exposure, we would expect the number of births at age 14 to increase by a factor 6/5; the number at age 13 to increase by a factor 5/3; and the number at age 12 to increase by a factor 3/1=3. The expected number of *additional births, compared with the original design*, would be these factors, minus 1, times the number of observed births: $(1/5)*b_{14}$; $(2/3)*b_{13}$; and $2*b_{12}$ for ages 14, 13, and 12, respectively.

Next consider a further reduction of the minimum age for eligibility to age 13. The line on the Lexis diagram that separates the blue and red areas would be lowered by one more year. There would then be a full 3 years of observation of age 14 and age 13, and 2.5 years of observation of age 12. The new factors for the original number of births would be 6/5 for age 14, 2 for age 13, and 5 for age 12. Subtracting 1 from each factor, the expected number of *additional births, compared with the original design,* would be $(1/5)*b_{14}$; b_{13}; and $4*b_{12}$ for ages 14, 13, and 12, respectively.

If the comparison in the preceding paragraph is revised so that the number of under-15 births in a survey with minimum age 13 is to be compared with the number in a survey with minimum age 14, then the expected number of additional births would be 0, $(1/3)*b_{13}$; and $2*b_{12}$ for ages 14, 13, and 12, respectively.

Table 3.2 The additional number of under-15 births in the past 3 years expected if a survey's minimum age for eligibility were reduced from age 15 to age 14 or age 13

		Simulation of minimum age 14				Simulation of minimum age 13			
Country	Observed number of under-15 births	Additional births	Additional respondents	Respondents per additional birth	Percent increase in number of births	Additional births	Additional respondents	Respondents per additional birth	Percent increase in number of births
Angola	81	23	894	39	28	6	889	148	36
Bangladesh	104	30	910	30	29	7	924	132	36
Chad	58	19	1323	70	33	5	1438	288	41
Colombia	36	10	1610	161	28	2	1508	754	33
Madagascar	55	22	1002	46	40	11	1205	110	60
Nigeria	65	18	1345	75	28	5	2172	434	35
Sierra Leone	45	17	633	37	38	7	980	140	53

Table 3.2 lists in alphabetical order the six countries/surveys for which the simulated number of additional births would be greatest—Angola, Bangladesh, Chad, Madagascar, Nigeria, and Sierra Leone—as well as Colombia, for which ages 13 and 14 were actually included in the survey. These are not necessarily the

countries with the highest under-15 fertility, because the number of births is roughly proportional to the sample size. For example, Colombia has low under-15 fertility, but its 2015 survey was large, so the number of additional cases would be relatively large. The first column gives the observed number of under-15 births in the past 3 years in these surveys. The next three columns refer to a reduction of the minimum age for eligibility to age 14, and the last three columns refer to a further reduction to age 13. The numbers of births are rounded weighted numbers. (The numbers of respondents are unweighted but closely match the weighted numbers because of how the weights in the household survey are normalized.) The third column in each group is the ratio of the number of additional respondents in the women's survey to the estimated number of additional births (i.e. cases/births).

For example, in the Angola survey a total of 81 under-15 births were observed for the past 3 years. If the minimum age for eligibility had been reduced to age 14, an additional 894 14-year-old *de facto* girls would have been included in the women's survey,[15] and we would expect them to have added 23 under-15 births in the last 3 years. An average of 39 interviews with these girls would have been required to obtain one additional under-15 birth. If the minimum age had been reduced further, to 13, the sample size for the women's survey would have been increased by an additional 889 13-year-old *de facto* girls, with an increase of six more under-15 births. An average of 148 interviews with 13-year-old girls would have been required for one additional under-15 birth.

Bangladesh is included in this table but is an exception because it was an EMW survey. An extension below age 15 would have to be limited to ever-married girls. In this survey, about one-fifth of the 15-year-old girls satisfied the EMW criterion. The proportion would be much lower for ages below 15, and the number of eligible girls would be small. Previous DHS surveys in Bangladesh (1993-94 through 2010) had minimum age 10 or (in 2011) minimum age 12. The minimum was 12 in an earlier survey in Turkey (1993) and 13 in an earlier survey in India (1992-93).

The Colombia 2015 survey is also exceptional. It is included in the table because in this survey, as mentioned, the minimum age of eligibility actually extended below age 15, to age 13. Nearly all surveys conducted in Colombia have been independent of USAID and DHS except for a minimal degree of standardization and coordination. Beginning with the survey conducted in 2004-05, Colombia has routinely included girls age 13 and 14 in its surveys of women. We believe that Colombia is the only country without an EMW criterion for eligibility that has obtained birth histories from girls below age 15. In the 2015 survey, the reduced age of eligibility brought in 1,332 additional girls age 13 and 1,407 girls age 14.

When the simulation procedure is applied to the Colombia survey, with girls age 13 and 14 omitted from the actual data but then simulating the effect of including them, we estimate that inclusion of the 14-year-olds would add ten under-15 births in the past 3 years, and the inclusion of the 13-year-olds would add two

[15] The simulated numbers of additional cases in the survey of women are not adjusted for possible refusals and non-response. They are also not adjusted for age displacement around age 15. In some surveys, interviewers will shift girls who are actually age 15 into age 14 in order to reduce their workload. In other surveys, the effort to avoid such transfers is too aggressive. For example, in the Sierra Leone 2013 household survey, included in the table, 633 girls are reported at age 14 and 1,063 at age 15. There was clearly age displacement from 14 to 15 in this survey, probably due to overreaction during training and supervision to transfers from age 15 to 14 in the 2008 survey, in which 988 girls were reported at age 14 and only 245 at age 15. It is also possible that girls who have a birth before age 15 will tend to be misreported at age 15 (or older) because of a bias in the direction of normative combinations of age and childbearing.

such births. These estimates are consistent with the actual data from the survey. The 14-year-old girls actually included in the survey of women added eight births; the 13-year-old girls added one birth. (These are weighted frequencies, rounded.) The expected frequencies (10 and 2) and the observed frequencies (8 and 1) are virtually indistinguishable, especially considering the size of the survey.

To describe the gain in precision that would result from an increased number of births, we go back to the observation in Section 2.8 that if the original total number of births is B, and the simulated number is $\hat{B}$, then the percentage reduction in the width of the confidence interval of a pooled rate would be (as a rough approximation) about $100 * [1 - sqrt\left(\frac{B}{\hat{B}}\right)]$. For example, if B increased by 30%, the width of the confidence interval would be reduced by about 12%. If B increased by 60%, the width would be reduced by about 21%. The ratios of additional interviews to additional births in Table 3.2, and the corresponding improvements in precision, make it unlikely that a reduction in the minimum age of eligibility would be cost-effective, at least in terms of improving the fertility estimates.

A final observation with regard to precision—that is, the widths of the confidence intervals—is that the intervals are usually, but not always, found to be narrower for the PR estimates than for the Lexis estimates or Equal estimates. The difference is not large enough to alter a conclusion that the alternative estimates are very similar in their levels and their precision.

4 DISCUSSION

This methodological report has focused narrowly on the question of how DHS surveys can be used to estimate under-15 fertility. We have shown that, despite the fact that age 15 is the lower end of the age range for eligibility in standard DHS surveys, the birth histories of women age 15-19 are an adequate source of data for estimates of recent under-15 fertility. A reduction in the minimum age to 14 or 13, for example, would yield more births but would not add substantially to the precision of the estimates.

This report is closely related to two other DHS reports that include the topic of under-15 fertility: Occasional Paper 9 (Way 2014), on issues in collecting data on adolescents; and Comparative Report 45 (MacQuarrie, Mallick, and Allen 2017), on adolescent sexual and reproductive health. All of these reports are a response to programmatic interest in averting early childbearing. The present methodological report includes estimates from 67 countries and surveys, with emphasis on measurement and interpretation.

The typical age-specific rates in DHS survey reports describe a 5-year age interval during the 3 years before the survey, or a 5-year age interval during the 5 years before the survey. DHS does not normally recommend the calculation of age-specific fertility rates for single years of age. In this report we do include rates for single years of age, within the 3 years or 5 years before the survey. Those rates are of interest in themselves and are the building blocks for constructing the pooled rates for wider age intervals. The three pooled rates are R1, the rate for age 12-14 during the 3 years before the survey; R2, the rate for age 10-14 during the 3 years before the survey; and R3, the rate for age 10-14 during the 5 years before the survey. R1 is introduced because age 12-14 is the interval for which we have data on the 3 years before the survey. Women age 15-19 at the time of the survey have no exposure to age 10-11 during the past 3 years. It is necessary to assume that there are no births at age 10-11 in order to extend the estimate to the full age range 10-14. Interest is primarily in R2 and secondarily in R3, but R2 is almost entirely determined by R1.

For each of these three rates, we describe three alternative ways to produce an estimate. Each is expressed as a different weighting of the underlying rates for single years of age. The first alternative, which makes the most complete use of the available data, brings in information about the number of girls in the age range 12-14 or 10-14 in the household survey. That information is in the PR data file, so it is convenient to refer to these as the PR estimates. A second set of estimates uses the geometry of a Lexis diagram to re-weight the single-year rates; these are the Lexis estimates. The third option simply takes the arithmetic average of the single-year rates; these are the Equal estimates.

The analysis showed that the three alternative versions of R1, R2, and R3, when estimated for the 67 different surveys and countries, agree closely. The estimates are virtually always within 1 point (one birth per 1,000 woman-years of exposure) of one another, or within 10% of one another, or within one another's 95% confidence intervals. When there is a noticeable difference, it appears to be attributable to variation in the age distribution of exposure within the age interval 10-14, which is in turn due to poor quality of reports of current age in the age interval 10-19. The Lexis and Equal approaches involve an assumption that the proportion of such exposure that is within the age range 12-14 is 0.60. The data suggest that in about a quarter of the surveys, the correct proportion is substantially lower, and the assumed value of 0.60 can induce an upward bias of 10% or more. The estimated distribution of exposure within age 12-14 can be uneven or implausible because of misreporting around age 15. Age displacement across the boundary of the 15th birthday is well documented. There are other common types of age misreporting within the full age range 10-19 (such as a preference for 12 and 18 and an avoidance of 19), and when ages are backdated by 3 or 5 years and combined to construct weights, these errors can produce unlikely distributions of exposure.

Beginning in 2017, DHS has provided the Lexis versions of R2 and R3 in the main reports and on STATcompiler. R3_Lexis was also the version used in Comparative Report 45. It has been adopted because of a strong demographic rationale. It also is considerably easier to calculate than the PR method. Computational simplicity is not a criterion for DHS itself, but it is relevant for the community of users who may wish to replicate DHS indicators. One motivation for this report was a need to confirm and document the Lexis estimates. The empirical comparison with the more complex PR method and the even simpler Equal method leads to a validation of the Lexis method, but when the three methods usually agree closely, the choice of method is not critical. The PR method makes the most complete use of the data, but it is also most sensitive to errors; it may be the best method if the data, specifically the reports of current age for ages 10-19, are of good quality, but by the same token it may be the worst method if the age reports are poor. The Equal method is surprisingly robust and is least affected by age errors. The Lexis method is intermediate in its sensitivity to data quality.

When the age reports are poor, the estimates of single-year rates will certainly be affected. This report has taken those rates at face value and used them to construct the estimates for age intervals, but a complete analysis of the sensitivity to reporting errors would include the effect of errors on the single-year rates.

We have shown how to simulate the number of additional under-15 births that would be expected if the minimum age for eligibility for a survey could be reduced to age 14 or 13 and how to estimate the improvement in precision of an under-15 fertility rate arising from inclusion of the additional births. It is unlikely that the additional improvement in precision would justify a reduction to 14 or 13, even apart from the other considerations that favor retaining the current minimum age.

Under-15 fertility is very low in almost all surveys. The under-15 rates described here are helpful for identifying countries with higher levels of under-15 fertility, but the simplest indicator of adolescent fertility is just the percentage of women age 15-19 at the time of a survey who had a birth before age 15. Also, as was shown in Comparative Report 45, there is a high correlation, at an aggregate level, between under-15 fertility and fertility at age 15-19. Typically, within the age range 10-19 the risk of childbearing increases monotonically from one year of age to the next. If fertility is relatively low at age 15-19, it is also relatively low before 15.

We recommend caution in the interpretation of under-15 fertility estimates that go back more than 5 years before the survey. It is possible to identify under-15 births in the birth histories from women of any age at the time of the survey, but it is likely that estimates for earlier time periods would tend to be biased because of potential displacement of current age or dates of birth to more normatively acceptable combinations. There is no reason to believe that reports of early births would be more accurate, or even equally accurate, when obtained from women age 20 and above. There is always more interest in the most recent time period, and trends can be inferred from the sequence of surveys available for most countries.

For surveys limited to ever-married women (EMW), fertility rates for age 15 and above require an adjustment with an "all-women" factor. The methods described here apply to EMW surveys, but will be harder to interpret.

DHS does not intend to incorporate under-15 fertility in the total fertility rate (TFR). Births under age 15 have always been included in the DHS version of the general fertility rate (GFR). Because of the small number of births before age 15, fertility rates for adolescents that are disaggregated by region, wealth quintile, etc., although available on STATcompiler, should be interpreted with caution. There will be little statistical power for inferring differences between subpopulations. Because of the high correspondence between fertility at age 10-14 and fertility at age 15-19, an analysis of differentials within age 15-19 should be an excellent guide to differentials at younger ages.

APPENDIX A

Appendix Table A1 List of the 67 surveys included in this analysis, with estimates of the fertility rates (births per 1,000 years of exposure) for single years of age

		3 years before the survey			5 years before the survey				
Country	Year	Age 12	Age 13	Age 14	Age 10	Age 11	Age 12	Age 13	Age 14
Afghanistan	2015	0.0000	0.0649	0.7808	0.0000	0.0000	0.2537	1.3938	2.9047
Albania	2008	0.0000	0.0000	1.2574	0.0000	0.0000	0.0000	0.0000	0.7660
Angola	2015	5.8418	6.6649	40.8689	0.0000	0.0000	3.6576	11.3450	38.5025
Armenia	2015	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Azerbaijan	2006	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Bangladesh	2014	0.0000	14.7237	36.3259	0.0000	0.0000	0.8153	10.9941	34.0536
Benin	2011	5.4094	1.7578	5.0992	0.0000	0.0000	3.5445	3.6579	9.2589
Bolivia	2008	0.0000	2.5128	11.1485	0.0000	0.5780	0.0000	2.0770	11.3441
Burkina Faso	2010	0.0000	0.0000	3.1161	0.0000	0.0000	0.0000	0.9580	5.3166
Burundi	2016	0.0000	1.3000	2.3454	0.0000	0.0000	0.0000	1.5220	2.1687
Cambodia	2014	0.0000	0.0000	0.5035	0.0000	0.0000	0.0000	0.0000	1.4928
Cameroon	2011	0.0000	9.6794	21.9162	0.0000	0.0000	5.2831	12.2445	22.1036
Chad	2014	0.0000	11.6987	20.3922	0.0000	0.0000	5.5330	13.3974	34.1659
Colombia	2015	0.3447	2.5128	9.7754	0.0000	0.0000	0.0811	2.3267	11.6405
Comoros	2012	0.0000	4.6258	4.0250	0.0000	0.0000	0.0000	1.9924	6.1388
Congo	2011	1.1190	3.2297	19.7805	0.0000	0.0000	1.2933	5.0727	18.5918
Congo Democratic Republic	2013	0.0000	6.9292	10.7362	0.0000	0.0000	1.8612	7.4328	18.0868
Côte d'Ivoire	2011	0.0000	2.6883	25.3482	0.0000	0.0000	7.4450	3.4870	24.8770
Dominican Republic	2013	0.0000	2.7727	9.5085	0.0000	0.0000	0.0000	7.6100	11.8309
Egypt	2014	0.0000	0.8202	1.4120	0.0000	0.0000	0.0000	1.0418	1.5401
Ethiopia	2016	0.0000	0.0000	3.3874	0.0000	0.0000	2.0173	0.0362	3.3371
Gabon	2012	0.0000	16.7012	15.0669	0.0000	0.0000	1.1491	10.8209	18.9512
Gambia	2013	0.0000	0.0000	9.8881	0.0000	0.0000	0.0000	0.5401	13.0466
Ghana	2014	0.0000	1.4373	4.3087	0.0000	0.0000	0.0000	0.7023	4.3712
Guatemala	2014	0.0000	1.8920	15.6520	0.0000	0.0000	0.0000	2.5564	14.1822
Guinea	2012	0.0000	18.6690	27.4983	0.0000	0.0000	3.3800	17.0862	32.2472
Guyana	2009	0.0000	2.1777	2.5726	0.0000	0.0000	0.0000	7.2365	8.5570
Haiti	2016	2.2360	2.8073	1.6683	0.0000	0.0000	0.4929	1.3983	1.9121
Honduras	2011	0.0000	2.5842	15.9662	0.0000	0.0000	0.2294	2.7137	16.0496
India	2015	0.0000	0.7066	1.7959	0.0000	0.0000	0.0000	0.6104	0.9121
Indonesia	2012	0.6746	0.0000	0.6444	0.0000	0.0000	0.2772	0.2708	1.4372
Jordan	2012	0.0000	0.0000	0.2537	0.0000	0.0000	0.0000	0.0000	0.1371
Kenya	2014	1.4198	0.0537	8.8344	0.0000	0.0000	0.3175	1.4381	9.1168
Kyrgyz Republic	2012	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Lesotho	2014	0.0000	0.0000	4.8496	0.0000	0.0000	0.0000	0.0000	3.4800
Liberia	2013	0.0000	2.9650	12.7781	0.0000	0.0000	1.5811	3.6484	21.1059
Madagascar	2008	9.0356	8.6522	20.5132	0.0000	0.0000	6.2928	15.0586	25.4336
Malawi	2015	3.3486	3.0656	8.9924	0.0000	0.0000	1.1905	2.6005	10.1141
Maldives	2009	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Mali	2012	0.0000	17.9397	37.0895	0.0000	0.0000	5.8984	17.2616	58.0118
Moldova	2005	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Morocco	2003	0.0000	0.0000	0.4890	0.0000	0.0000	0.0000	0.0000	1.2014
Mozambique	2011	0.0000	5.8954	24.2264	0.0000	0.0000	0.7477	6.0246	25.3866
Myanmar	2015	0.0000	2.9348	0.9137	0.0000	0.0000	0.0000	2.5350	0.5097
Namibia	2013	2.1665	15.1168	2.5111	0.0000	0.0000	0.4530	6.5449	4.3581
Nepal	2016	0.0000	0.0000	2.5294	0.0000	0.0000	0.0000	0.7932	6.1021
Nicaragua	2001	0.0000	2.2891	20.0976	0.0000	0.0000	0.1779	2.2144	19.2561
Niger	2012	0.6682	3.6092	25.0829	0.0000	0.0000	5.7160	7.5696	31.6828
Nigeria	2013	0.8002	2.8732	13.1568	0.0000	0.0000	1.3831	5.5587	15.8441
Pakistan	2012	0.0000	0.0000	0.1361	0.0000	0.0000	0.9967	0.3284	0.3073
Peru	2012	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Philippines	2013	0.0000	0.0000	4.1400	0.0000	0.0000	0.0000	0.0000	3.4536
Rwanda	2014	1.0964	0.0000	2.8390	0.0000	0.0000	0.2887	0.0000	1.6511
São Tomé and Principe	2008	0.0000	0.0000	8.0766	0.0000	0.0000	0.0000	0.0000	8.6490
Senegal	2016	0.0000	1.5699	6.3248	0.0000	0.0000	0.0000	3.3957	5.5794
Sierra Leone	2013	3.5035	5.6253	16.6906	0.0000	0.8698	2.2414	6.5369	23.9124
Swaziland	2006	0.0000	5.0440	12.5922	0.0000	0.0000	0.0000	3.0457	9.7060
Tajikistan	2012	0.0000	0.0000	1.0800	0.0000	0.0000	0.0000	0.0000	0.6019
Tanzania	2015	0.0000	0.7724	2.4065	0.0000	0.0000	1.2387	0.3659	2.9151
Timor-Leste	2016	0.0000	0.0000	1.8563	0.0000	0.0000	0.0000	0.0000	2.3377
Togo	2013	0.0000	6.3779	4.4412	0.0000	0.0000	2.0461	5.3306	6.4026
Turkey	2003	0.0000	0.0000	0.8197	0.0000	0.0000	1.0728	1.0392	0.7747
Uganda	2016	2.1226	0.5396	7.4204	0.0000	0.0000	1.9222	2.7039	7.6929
Ukraine	2007	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.2062
Yemen	2013	0.0000	0.7127	4.6630	0.0000	0.0000	0.0000	1.1205	5.0983
Zambia	2013	0.0000	0.0000	10.6496	0.0000	0.0000	1.3259	0.3801	10.3776
Zimbabwe	2015	0.0000	0.0000	8.9006	0.0000	0.0000	0.0000	0.0000	5.3250

Appendix Table A2 Estimates of rates R1, R2, and R3 using the PR, Lexis, and "Equal" methods, and the proportion P of girls age 10-14 who are 12-14, estimated from the household data

		Estimates of R1			Estimates of R2			Estimates of R3			
Country	Survey year	PR	Lexis	Equal	PR	Lexis	Equal	PR	Lexis	Equal	P
Afghanistan	2015	0.2753	0.2911	0.2819	0.1603	0.1746	0.1691	0.8873	0.9306	0.9104	0.5822
Albania	2008	0.4227	0.4164	0.4191	0.2580	0.2498	0.2515	0.1467	0.1425	0.1532	0.6104
Angola	2015	16.4686	17.7720	17.7919	9.0812	10.6632	10.6751	9.3779	10.6398	10.7010	0.5514
Armenia	2015	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.5843
Azerbaijan	2006	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.6014
Bangladesh	2014	16.9913	16.6597	17.0165	10.1711	9.9958	10.2099	9.1676	8.9676	9.1726	0.5986
Benin	2011	4.1199	4.0633	4.0888	2.1236	2.4380	2.4533	2.6906	3.2311	3.2923	0.5155
Bolivia	2008	4.1201	4.5764	4.5538	2.3056	2.7458	2.7323	2.4530	2.7775	2.7998	0.5596
Burkina Faso	2010	0.8499	1.0071	1.0387	0.4442	0.6043	0.6232	0.9944	1.2172	1.2549	0.5226
Burundi	2016	1.1367	1.2151	1.2151	0.6137	0.7291	0.7291	0.6422	0.7126	0.7381	0.5399
Cambodia	2014	0.1491	0.1547	0.1678	0.0849	0.0928	0.1007	0.2593	0.2782	0.2986	0.5692
Cameroon	2011	10.0415	10.0885	10.5319	5.5810	6.0531	6.3191	7.4833	7.7502	7.9262	0.5558
Chad	2014	9.3235	9.9822	10.6970	4.9680	5.9893	6.4182	8.6469	9.8575	10.6192	0.5329
Colombia	2015	4.0943	4.1571	4.2110	2.4323	2.4943	2.5266	2.6890	2.7584	2.8097	0.5941
Comoros	2012	2.8562	2.8792	2.8836	1.6881	1.7275	1.7302	1.6054	1.6255	1.6262	0.5910
Congo	2011	7.7178	7.7200	8.0431	4.4237	4.6320	4.8258	4.6187	4.7425	4.9915	0.5732
Congo Democratic Republic	2013	5.2531	5.9040	5.8885	2.8495	3.5424	3.5331	4.6264	5.5208	5.4762	0.5424
Côte d'Ivoire	2011	8.1496	8.9133	9.3455	4.4558	5.3480	5.6073	6.3193	7.0258	7.1618	0.5467
Dominican Republic	2013	3.8480	4.0235	4.0937	2.2051	2.4141	2.4562	3.6550	3.8877	3.8882	0.5730
Egypt	2014	0.7373	0.7733	0.7441	0.4272	0.4640	0.4464	0.5060	0.5337	0.5164	0.5794
Ethiopia	2016	0.9486	1.0448	1.1291	0.5073	0.6269	0.6775	0.9509	1.0168	1.0781	0.5348
Gabon	2012	9.9434	10.8750	10.5894	5.5838	6.5250	6.3536	5.5658	6.4823	6.1843	0.5616
Gambia	2013	2.8648	3.1793	3.2960	1.6263	1.9076	1.9776	2.4633	2.7971	2.7173	0.5677
Ghana	2014	1.7004	1.7916	1.9153	0.9793	1.0750	1.1492	0.8303	0.9127	1.0147	0.5759
Guatemala	2014	5.7604	5.7619	5.8480	3.4315	3.4571	3.5088	3.2011	3.1979	3.3477	0.5957
Guinea	2012	13.5619	14.2104	15.3891	7.2289	8.5262	9.2335	8.9551	10.0958	10.5427	0.5330
Guyana	2009	1.5350	1.5920	1.5834	0.8810	0.9552	0.9500	2.9078	3.0511	3.1587	0.5739
Haiti	2016	2.2252	2.2352	2.2372	1.3371	1.3411	1.3423	0.7717	0.7374	0.7606	0.6009
Honduras	2011	5.9900	6.2818	6.1835	3.4523	3.7691	3.7101	3.5285	3.7243	3.7985	0.5763
India	2015	0.4197	0.8385	0.8342	0.1779	0.5031	0.5005	0.1063	0.3157	0.3045	0.4238
Indonesia	2012	0.4408	0.4412	0.4397	0.2548	0.2647	0.2638	0.3727	0.3845	0.3970	0.5780
Jordan	2012	0.0892	0.0923	0.0846	0.0530	0.0554	0.0507	0.0289	0.0293	0.0274	0.5937
Kenya	2014	3.0954	3.2912	3.4360	1.6830	1.9747	2.0616	1.8224	2.0670	2.1745	0.5437
Kyrgyz Republic	2012	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.6090
Lesotho	2014	1.4285	1.5854	1.6165	0.7992	0.9512	0.9699	0.5827	0.6648	0.6960	0.5595
Liberia	2013	4.8581	5.0011	5.2477	2.6776	3.0007	3.1486	4.5359	4.7789	5.2671	0.5511
Madagascar	2008	12.3166	12.7688	12.7337	6.7355	7.6613	7.6402	8.4138	9.4562	9.3570	0.5469
Malawi	2015	4.8108	4.9787	5.1355	2.4870	2.9872	3.0813	2.2398	2.6144	2.7810	0.5169
Maldives	2009	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.4492
Mali	2012	15.0604	19.0557	18.3431	7.8585	11.4334	11.0058	12.4864	17.1067	16.2344	0.5218
Moldova	2005	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.6204
Morocco	2003	0.1617	0.1730	0.1630	0.0957	0.1038	0.0978	0.2331	0.2447	0.2403	0.5916
Mozambique	2011	9.5857	9.1978	10.0406	5.2245	5.5187	6.0244	5.6503	5.7504	6.4318	0.5450
Myanmar	2015	1.2393	1.3257	1.2828	0.6844	0.7954	0.7697	0.5524	0.6188	0.6089	0.5522
Namibia	2013	6.6532	6.5834	6.5981	3.9148	3.9500	3.9589	2.1342	2.2644	2.2712	0.5884
Nepal	2016	0.8100	0.8527	0.8431	0.4737	0.5116	0.5059	1.3155	1.3711	1.3791	0.5848
Nicaragua	2001	7.0249	7.3923	7.4623	3.9972	4.4354	4.4774	3.9656	4.2750	4.3297	0.5690
Niger	2012	7.7266	9.4769	9.7868	3.7153	5.6862	5.8721	6.7152	8.8500	8.9937	0.4808
Nigeria	2013	5.2516	5.0336	5.6101	2.9021	3.0202	3.3661	4.0862	4.0706	4.5572	0.5526
Pakistan	2012	0.0427	0.0546	0.0454	0.0241	0.0327	0.0272	0.3203	0.3611	0.3265	0.5635
Peru	2012	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.4284
Philippines	2013	1.2920	1.2904	1.3800	0.7619	0.7742	0.8280	0.6257	0.6323	0.6907	0.5897
Rwanda	2014	1.2806	1.2592	1.3118	0.7182	0.7555	0.7871	0.3520	0.3500	0.3880	0.5609
São Tomé and Principe	2008	2.2163	2.7350	2.6922	1.1958	1.6410	1.6153	1.4161	1.7824	1.7298	0.5395
Senegal	2016	2.4865	2.6454	2.6316	1.4127	1.5872	1.5789	1.6969	1.8339	1.7950	0.5681
Sierra Leone	2013	8.4228	7.7708	8.6065	4.9110	4.6625	5.1639	6.4481	5.8546	6.7121	0.5831
Swaziland	2006	5.5218	5.8230	5.8787	3.2063	3.4938	3.5272	2.3024	2.5058	2.5503	0.5807
Tajikistan	2012	0.3578	0.3864	0.3600	0.2128	0.2318	0.2160	0.1180	0.1251	0.1204	0.5949
Tanzania	2015	0.9593	0.9842	1.0596	0.5242	0.5905	0.6358	0.8110	0.8521	0.9039	0.5464
Timor-Leste	2016	0.5590	0.6065	0.6188	0.3247	0.3639	0.3713	0.3978	0.4358	0.4675	0.5809
Togo	2013	3.2829	3.4910	3.6064	1.7618	2.0946	2.1638	2.3681	2.7011	2.7559	0.5367
Turkey	2003	0.2674	0.2743	0.2732	0.1607	0.1646	0.1639	0.5754	0.5792	0.5774	0.6009
Uganda	2016	3.1889	3.3306	3.3609	1.7459	1.9984	2.0165	2.1725	2.3814	2.4638	0.5475
Ukraine	2007	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.2286	0.2396	0.2412	0.6053
Yemen	2013	1.6641	1.6581	1.7919	0.9628	0.9949	1.0751	1.1500	1.1750	1.2438	0.5786
Zambia	2013	2.8832	3.4894	3.5499	1.5858	2.0937	2.1299	1.9819	2.3741	2.4167	0.5500
Zimbabwe	2015	2.8182	2.8986	2.9669	1.6440	1.7392	1.7801	0.9575	0.9978	1.0650	0.5834

REFERENCES

MacQuarrie, K. L. D., L. Mallick, and C. Allen. 2017. *Sexual and Reproductive Health in Early and Later Adolescence: DHS Data on Youth Age 10-19.* DHS Comparative Reports 45. Rockville, Maryland: ICF. https://www.dhsprogram.com/pubs/pdf/CR45/CR45.pdf.

Pullum, T. W. 2004. "Natality: Measures Based on Censuses and Surveys Chapter 16 of *The Methods and Materials of Demography, 2nd edition,* pp. 407-428. D. A. Swanson and J. Siegel (eds.). London, England: Elsevier Academic Press.

Pullum, T. W., and S. Assaf. 2016. *Long Term Trajectories of Fertility and Contraceptive Use.* DHS Analytical Studies 58. Rockville, Maryland, USA: ICF. https://www.dhsprogram.com/pubs/pdf/AS58/AS58.pdf.

Pullum, T. W., S. Assaf, and S. Staveteig. 2017. *Comparisons of DHS Estimates of Fertility and Mortality with Other Estimates.* DHS Methodological Reports 21. Rockville, Maryland, USA: ICF. https://www.dhsprogram.com/pubs/pdf/MR21/MR21.pdf.

Pullum, T. W., and Sarah Staveteig. 2017. *An Assessment of the Quality and Consistency of Age and Date Reporting in DHS Surveys, 2000-2015.* DHS Methodological Reports No. 19. Rockville, Maryland, USA: ICF.

Rutstein, S. O., and G. Rojas. 2012. *Guide to DHS Statistics.* Calverton, Maryland, USA: ICF. https://www.dhsprogram.com/pubs/pdf/DHSG1/Guide_to_DHS_Statistics_29Oct2012_DHSG1.pdf.

Schoumaker, B. 2014. *Quality and Consistency of DHS Fertility Estimates, 1990 to 2012.* DHS Methodological Reports 12. Rockville, Maryland, USA: ICF. https://www.dhsprogram.com/pubs/pdf/MR12/MR12.pdf.

Shryock, H., J. Siegel, and Associates. 1971. *The Methods and Materials of Demography.* Washington, D.C.: U.S. Census Bureau/U.S. Government Printing Office.

Way, A. 2014. *Youth Data Collection in DHS Surveys: An Overview.* DHS Occasional Paper 9. Rockville, Maryland: ICF. https://www.dhsprogram.com/pubs/pdf/OP9/OP9.pdf.

WHO. 2018. *Guidance on Ethical Considerations in Planning and Reviewing Research Studies on Sexual and Reproductive Health in Adolescents.* Geneva, Switzerland: World Health Organization. http://apps.who.int/iris/bitstream/handle/10665/273792/9789241508414-eng.pdf?ua=1.