USAID
FROM THE AMERICAN PEOPLE

# COMPARISONS OF DHS ESTIMATES OF FERTILITY AND MORTALITY WITH OTHER ESTIMATES

# DHS METHODOLOGICAL REPORTS 21

AUGUST 2017

This publication was produced for review by the United States Agency for International Development. It was prepared by Thomas Pullum, Shireen Assaf, and Sarah Staveteig.

DHS Methodological Reports No. 21

# Comparisons of DHS Estimates of Fertility and Mortality with Other Estimates

Thomas Pullum

Shireen Assaf

Sarah Staveteig[1]

ICF

Rockville, Maryland, USA

August 2017

[1] Avenir Health

*Corresponding author:* Thomas 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:** Thanks to Amy Tsui, Linnea Zimmerman, Qingfeng Li, and Yoonjoung Choi at Johns Hopkins Univeristy and Yared Mekonnen at Mela Research for answering questions about PMA2020 data and fertility estimation methods, and to Mahmoud Elkasabi and Trevor Croft for helpful discussions of Stata sampling.

Editor: Diane Stoy
Document Production: Christopher Gramer

This study was carried out with support provided by the United States Agency for International Development (USAID) through The DHS Program (#GPO–C–00–08–00008–00). The views expressed are those of the authors and do not necessarily reflect the views of USAID or the United States Government.

The DHS Program assists countries worldwide in the collection and use of data to monitor and evaluate population, health, and nutrition programs. For additional information about The DHS Program, contact: The DHS Program, ICF, 530 Gaither Road, Suite 500, Rockville, MD 20850, USA; phone: +1 301-407-6500; fax: +1 301-407-6501; email: reports@dhsprogram.com; Internet: www.dhsprogram.com.

Recommended citation:

Pullum, Thomas, Shireen Assaf, and Sarah Staveteig. 2017. *Comparisons of DHS Estimates of Fertility and Mortality with Other Estimates*. DHS Methodological Reports No. 21. 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-level 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
Director, The DHS Program

# Abstract

DHS surveys provide many indicators that are used for program planning and monitoring. Some of these are particularly sensitive and attract attention when the results of a new survey are released. This methodological report focuses on six demographic indicators of widespread interest: the Total Fertility Rate (TFR), Infant Mortality Rate (IMR), Under-5 Mortality Rate (U5MR), the Adult Female Mortality Probability (AFMP), the Adult Male Mortality Probability (AMMP) and the Maternal Mortality Ratio (MMR). It is not unusual for the results of a new survey to be questioned because they differ from another source. The goal of this report is to provide guidance on how to determine whether the estimates of these important indicators are plausible and consistent with other sources, or not plausible.

Determining whether a DHS estimate is consistent with other sources is usually a matter of degree. Some differences are expected for a variety of reasons. The report includes a discussion of potential reasons for discrepancies. The report then analyzes 51 surveys conducted since 2010. The DHS estimates are systematically compared with estimates of the TFR, AFMP, and AMMP from the UN Population Division, estimates of the IMR and U5MR from the UN's Inter Agency Group for Child Mortality Estimation (IGME), and MMR estimates from the World Health Organization (WHO). The structure of the comparisons can be applied to other sources, although it is always necessary to account for the potential kinds of differences discussed earlier. Two specific surveys provide focus for the comparisons.

The report also illustrates a strategy for comparing fertility rates computed from DHS data with those from other sources. We compare DHS fertility rate estimates with those from Performance Monitoring and Accountability 2020 (PMA2020) surveys, explore the effect of a slightly different methodology for computing rates on DHS estimates, and simulate the effects of a different sampling strategy and rate estimation method on fertility rates. Using a standard methodology, fertility rates computed from PMA2020 data in five countries showed a range of 5% to 22% difference in TFRs and a 4% to 17% difference in adolescent fertility rates compared to results from DHS surveys conducted within a three-year timespan.

To assess the effect of alternate measurement on fertility rates from the exact same survey, we used 256 DHS datasets to compare the results of a 2-year 2-birth adjusted estimation technique versus a 3-year n-birth technique with the same data. Our results show an average of only about one percentage point difference in total fertility and adolescent fertility rates from the same data using the alternate technique. In addition to measurement differences, it is important to consider design effects of a different cluster sample. In Ethiopia, we simulated subsamples of DHS data with a cluster distribution similar to a corresponding PMA2020 survey and recomputed fertility rates with a 2-year 2-birth adjusted estimation technique. These combined differences in sampling and methodology produced an average of a 3% to 4% difference in rate estimates but could plausibly produce as much as a 10% difference in TFR and a 23% difference in adolescent fertility rates. Notably, since the PMA2020 cluster sample sizes are larger than those used by DHS, this simulation is likely an overestimate of the effect of sampling differences.

KEY WORDS: Fertility, PMA2020, under-5 mortality, adult mortality, maternal mortality, sample design, data quality

# 1. Introduction

Surveys conducted by The Demographic and Health Surveys Program (DHS) are a source of widely used estimates of current fertility, the mortality of children and adults, and maternal mortality. The estimates are published in the main reports of each survey, are accessible on STATcompiler (DHS Program 2017), and have been analyzed in many DHS publications. The estimates are used by USAID, UNICEF, WHO, and other agencies to assess the effectiveness of programs, and by the United Nations Population Division to estimate long-term trajectories of vital rates and population growth in many developing countries.

DHS frequently assesses the quality of these estimates. Recent DHS Methodological Reports have focused on the birth histories, which include information about births and child deaths (Pullum and Becker 2014), estimates of fertility (Schoumaker 2014), and estimates of maternal mortality (Ahmed et al. 2014). Another recent report (Pullum and Staveteig 2017) assesses the quality of reports of ages and dates, which are crucial for all rates. These assessments have concluded that most surveys have produced estimates of good quality, although a few surveys—often the same surveys for multiple indicators—have been of lower quality. Respondents' incomplete knowledge of age and the dates of vital events is a severe limitation in some contexts, but is certainly not the sole source of deficient data.

There have been a number of apparent inconsistencies between successive DHS surveys, or between DHS surveys and other sources. There have also been instances in which a DHS estimate was substantially higher or lower than expected, regardless of a specific comparison. This report will review some strategies for identifying and understanding such inconsistencies.

Non-DHS estimates are often based in part on DHS estimates, or derived from procedures that are similar to DHS procedures, and cannot be regarded as independent of DHS estimates. It is possible, for example, that DHS estimates of the IMR in a particular country are chronically low, because some children who died at a very early age were omitted from the birth histories. The estimates produced by the UN agencies, which rely to a large extent on DHS data, may have been adjusted upward in an attempt to compensate for those omissions. However, if they were not adjusted upward sufficiently, the series of adjusted estimates and projections may also be low. When the DHS estimates are compared with the UN estimates, the amount of bias in the DHS data will be under-estimated, or perhaps not identified.

Previous research on differences between estimates has concluded that discrepancies are due to differences in the data such as full versus abbreviated birth histories, or to differences in the statistical methods (Alkema et al. 2012, Brady and Hill 2017, Hancioglu and Arnold 2013, Masquelier et al. 2014, Silva 2012).

If a specific estimate is suspicious because it is not consistent with other sources, the analyst cannot conclude that it is erroneous. It is possible that some outcomes change abruptly, either temporarily or as part of a longer-term trend. However, it is always advisable to include more in-depth analysis in such a situation. Strategies for in-depth analysis of an estimate for a specific survey will not be described here. The purpose of this report is to provide guidance for judging whether in-depth internal checking, beyond what is normal for DHS surveys, is needed.

The report will also not review methods for adjusting data, since DHS has a policy of not adjusting estimates. There are several reasons for this policy. First, the main reports on DHS surveys are very comprehensive, with a wide variety of demographic and health indicators, and their publication cannot be delayed by a thorough analysis. Second, the data files are available at the same time as the main report is released, and these must be consistent with the report so that users can replicate and extend the information from the main report. If the national IMR were adjusted upwards, for example, then the microdata would not be consistent with the adjusted rate. Third, users may choose to make adjustments, and the DHS estimates in the reports that are consistent with the data files serve as a continuous reference or baseline against which a variety of alternative adjustments can be made. Thus, decisions about whether and how to adjust an estimate are made by the users, which include agencies such as the UN Population Division.

This report is intended primarily for users of the main reports that are published after each survey and for users of STATcompiler. For the benefit of users of DHS microdata files, some details will describe a specific statistical package, Stata, although they can be applied in other packages, such as SPSS, SAS, and R. These details are relevant to users of the data files and can be bypassed by other readers.

Data and methods will be described in the relevant chapters.

# 2. Potential Sources of Differences

This chapter will provide a brief review of reasons why two indicators from different sources may unexpectedly differ, but not actually be incompatible. The chapter serves as a checklist of potential issues that may be the source of the apparent difference. A more in-depth analysis is required when a difference remains even after the effect of these issues has been considered.

## 2.1. Statistical Issues

Many of the apparent discrepancies in DHS estimates can be attributed to sampling error in the DHS estimate, or sampling error or uncertainty in the number to which the DHS estimate is being compared. We will briefly review the statistical properties of DHS estimates, including how to interpret the standard error of an estimate, the relevance of denominators and numerators, the role of the sample design and how to take it into account, and the interpretation of confidence intervals, and how they differ from tests.

### *The standard error of an estimate*

Virtually every number in DHS reports is either a description of the sample—for example, the sample size—or an estimate of a population quantity, such as a mean or median, a proportion or percentage, or a rate or ratio. Sample estimates of a population quantity, or parameter, are statistics, and are subject to sampling variability.

If we were to draw repeated samples from the population, using the same sampling frame, sampling design, and sample size, the estimates of the same parameter would vary from one sample to another. The sampling error for a specific estimate is the deviation of that estimate from the "true" value in the population, or the statistic minus the parameter. Because we do not know the value of the parameter, the sampling error for the sample is not known. If the statistic is unbiased, then (by definition) the *expected* sampling error is zero and the average deviation over all possible repeated samples is zero.

The standard error can be interpreted as approximately the average of the *absolute values* of these deviations, over all possible samples. Statistical theory enables us to estimate the standard error, although we have only one sample and we do not know the true value of the parameter. The reports of DHS surveys include an Appendix B, which provides the estimated standard error for about 50 indicators in the surveys of women and of men, nationally and within the sampling strata. Standard errors are not included in STATcompiler, although a confidence interval for the MMR is included.

### *The role of the sample design*

With few exceptions, DHS samples are obtained in the following steps—skipping over the role of survey-specific considerations such as the budget and the desired statistical power.

The DHS samples are stratified, with strata generally corresponding to the urban and rural areas of each region (the level 1 administrative unit). The sample must satisfy some minimal criteria, which is specified in terms of the anticipated standard errors within each stratum. Strata with relatively small populations tend to be over-sampled, while strata with relatively large populations are often under-sampled. The sample thus has a distribution that is somewhat more equally allocated across strata than the population distribution, in order to optimize the standard errors of within-stratum estimates. The use of strata and different sampling fractions for the strata requires adjustments during analysis with sampling weights.

The sampling frame is generally a list of all enumeration areas (EAs) in the most recent census, with an estimate of the number of households in each EA. Within each stratum, the EAs are sampled with probability proportional to size, where "size" is measured by the number of households in the sampling frame. The constant of proportionality in each stratum is determined by the desired sample size within that stratum. The EAs are also described as primary sampling units (PSUs) or clusters. Within each selected cluster, a constant number of households (such as 30) are selected during the household listing.

Households within the same cluster tend to resemble one another more closely than households selected at random from different clusters. This overlap of information from the same cluster requires other adjustments to the standard errors of stratum-level estimates or national-level estimates during analysis. Adjustments to the weights are also required because of the variation in nonresponse.

DHS recommends three types of adjustments during the analysis. The first is the use of sample weights that are required to produce unbiased estimates of population quantities. Weights can also alter the standard errors. When weights are used, the standard errors tend to increase. The second adjustment is for clustering. If this adjustment is not made, the estimates of the standard errors will be too small. That is, if the cluster adjustment is used, the standard errors tend to increase. The third adjustment is for stratification. If this adjustment is not made, the estimates of the standard errors will be too large. The adjustment tends to reduce them. The adjustments for clustering and stratification do not alter the point estimates, only the standard errors.[1]

The typical DHS sample design reduces the cost of fieldwork, relative to a simple random sample, but it also effectively reduces the sample size. Appendix B of the survey reports provides the actual sample size, or denominator, for each indicator and the design effect (DEFT). The design effect can be interpreted as the ratio of the actual sample size to the size of an equivalent simple random sample. If, for example, N=10,000 and DEFT=2, then the adjusted standard error will be equivalent to the unadjusted standard error of a sample of size 10,000/2 = 5,000 cases. The DEFT is virtually always greater than 1, although it can be less than 1 if the increase in efficiency due to stratification has more impact than the reduction of efficiency due to clustering. Adjusting the standard errors is very relevant to this report, because a difference that is statistically significant without the adjustment can easily become insignificant with the adjustment, which increases the standard error.

### *The relevance of denominators and numerators*

The sample size for an indicator, typically the denominator, is very important for conclusions about statistical significance. Appendix B of the survey reports makes this clear. In a simple random sample, a standard error is inversely proportional to the square root of the sample size. In a more complex sample, a standard error is inversely proportional to the square root of the effective sample size. A chi-square test statistic is directly proportional to the effective sample size. And, a $z$ or $t$ test statistic is directly proportional to the square root of the effective sample size. For rare events, the number of cases in the numerator is also important. For example, an infant mortality rate can be written as $r=d/b$, where $b$ is the size of a cohort of births and $d$ is the number of deaths to that cohort before the first birthday. That is, $d$ is the numerator and $b$ is the denominator. For a simple random sample, $d$ will have a Poisson sampling distribution and the standard error of the rate is approximately [sqrt($d$)]/$b$. This estimate can be written in two different ways. The first is [sqrt($r$)]/[sqrt($b$)], which confirms that the standard error is inversely proportional to the square

[1] Researchers using Stata can make these adjustments with the svyset and svy commands. All the estimates in the reports, including Appendix B, and on STATcompiler, have made the same adjustments but with different software, so they may differ slightly from the estimates from Stata or another package. The adjustments currently possible with DHS data are incomplete. Ideally, DHS datasets would provide two weights, one that is the inverse of the sampling fraction for the selection of clusters within strata, and one that is the inverse of the sampling fraction for the selection of households with clusters. It is not possible for DHS to separate these two components, because knowledge of the cluster-level weight could enable the identification of the actual cluster by anyone with access to the sampling frame. The weight variable included in DHS data files is the normalized product of those two weights.

root of the denominator. Alternatively, the standard error is $r$/[sqrt($d$)], which shows that the standard error is also inversely proportional to the square root of the numerator. The second relationship is particularly relevant for analysis of the MMR, because even in large surveys, the number of maternal deaths to sisters in the standard time interval (the 7 years before the survey) is often quite small and less than 100.

***Tests of differences***

Many statistical tests of hypotheses about DHS rates can be constructed without using the data files, if the standard error can be located in STATcompiler or in Appendix B of the reports. These include tests of differences between surveys, including two surveys conducted in the same country, or differences from specific values from another source. Some estimates from other sources, such as the UN Population Division, are not accompanied by standard errors or confidence intervals.[2] Tests can be one-tailed or two-tailed.

Many users of DHS data make a judgment about whether a difference between two estimates is significant by checking whether the 95% confidence intervals for the two estimates overlap. This technique is not equivalent to a test, except when there is a good separation between the confidence intervals. In that case, the user lacks a specific p-value for the test. Another weakness of the technique is that there is often a clear one-sided nature to the research hypothesis—for example, that the MMR has declined between two surveys, whereas confidence intervals are always two-sided.

To illustrate how to test whether a change in the MMR from one survey to the next was statistically significant, we will use two successive surveys from Zimbabwe, conducted in 2010 and 2015. In this example, the confidence intervals around the two estimates overlap, but there was a highly significant decline. An inference that there was not a significant decline, based on the overlap of the confidence intervals, would have been incorrect.[3]

If we go to STATcompiler and download the MMR estimates for the two surveys,[4] we find estimates of 960 and 651 maternal deaths per 100,000 births, respectively, for the 2010 and 2015 surveys. These numbers imply that there was a substantial decline from the first survey to the second, but was it statistically significant and can it be interpreted as a real change in the population?

STATcompiler provides confidence intervals to accompany the point estimates of the MMR. In the first survey, the interval extends from 778 to 1142. The second interval extends from 473 to 829. Each interval was calculated by adding and subtracting 2*se to each point estimate, where "se" is the standard error.

Testing is conceptually different from estimation, although both involve the estimate and the standard error of the estimate. The information required to conduct the relevant tests is contained in the estimates just given. We can conduct the test of a difference with five steps. Refer to the two surveys as #1 and #2, in chronological order. The goal is to determine the significance of the change from #1 to #2.

Step 1. Find the standard error (se) of each MMR estimate by dividing the width of each confidence interval by 4. Thus, the se for the 2010 estimate is (1142-778)/4= 91 and the se for the 2015 estimate is (829-473)/4 = 89.

---

[2] Uncertainty intervals, which may be provided by sources other than DHS, are not equivalent to confidence intervals.
[3] There is always a probability (given by the $p$-value) that any inference from a test is incorrect.
[4] There is approximately a two-year overlap in the reference periods, because the reference period is 7 years and the surveys are 5 years apart. The only way to avoid overlap would be to re-estimate with the data files. The effect of overlap is to bias the test toward the null hypothesis of no change.

Step 2. Calculate the differences between the pairs of estimates by subtracting the earlier estimate from the later estimate. If there was a decline over time, the difference is negative. If there was an increase over time, the difference is positive. The difference is 651-960=-309.

Step 3. Calculate the standard error of the difference. This is a simple calculation because the successive surveys are statistically independent. The standard error of the difference calculated in Step 2 is just the square root of the sum of the squares of the standard errors from the two surveys. Here, the standard error of the difference between surveys #1 and #2 is the square root of the sum of 91 squared and 89 squared. This is 127.29. (Rounding is deferred until the end of the calculations.)

Step 4. Divide the difference by its standard error. This will produce a $z$ test statistic, which will have an approximately normal sampling distribution with expected value 0 and standard deviation 1 under the null hypothesis that there was no change in the population MMR. The $z$ score is -309/127.29 =-2.43.

Step 5. Compare the test statistics with the critical values of the $z$ distribution. The standard criterion is the .05 level of significance. If the research hypothesis is that the population MMR changed between the two surveys, moving either up or down, and the null hypothesis is that it did not change, then the null hypothesis will be rejected if z is outside the range from -1.96 to +1.96. Clearly, the calculated z value is outside this range. We reject the null hypothesis that there was no change.

The research hypothesis, however, is actually that the MMR declined between two surveys. That is, there is a direction to our expectation for change in the MMR, and the test should be one-sided rather than two-sided. The critical value for this test is -1.65. If the calculated $z$ value is less than -1.65, then we infer that the decline was significant with $p<.05$. If z is less than -2.33, then we infer that the decline was significant with $p<.01$. The calculated $z$ value, -2.43, is therefore highly significant. The $p$-value is .008.

This approach is more accurate than checking whether the two confidence intervals overlap but is itself a shortcut. A more accurate approach would require going to the data files and using a statistical model with a logit link function and adjustments for clusters and strata, as well as sampling weights, in both surveys. If the approach described here produces a borderline result, then it would be advisable to use the more accurate approach.

### 2.2. The Reference Period

DHS estimates of fertility and mortality are always calculated for a reference period of time, such as the 3, 5, or 7 years before the survey, because they come from the retrospective birth histories or sibling histories. In the main reports, some national estimates are published with one reference period and another estimate for subpopulations. With adjustments to software, it is possible to construct estimates for intervals expressed in calendar years, rather than years ago, but these estimates must refer to an interval of time.

A common misinterpretation of the rates from DHS surveys is that they refer to the date of the survey. For example, the estimates of the MMR from the 2010 and 2015 surveys of Zimbabwe could be misinterpreted as applying to 2010 and 2015. They actually refer to the 7 years prior to those surveys—an interval of 1 to 84 months prior to the month of interview for each woman in the survey. The best choice of a time point would be the midpoint of that time interval, which would be approximately 3.5 years before the mean date of interview.

Estimates from UN agencies may refer to an interval of time, expressed as calendar years, or to points in time. For example, WHO provides estimates for 2015 that can be interpreted as a point in time, specifically the midpoint of 2015 or July 1, 2015. This manner of dating the estimates is possible because WHO applies statistical models for the MMR, which synthesize many data sources (including DHS) with varying intervals for each source. The models fit a continuous trend line through these sources, and the points on that line can be interpreted as estimates for specific time points. The best comparison with a DHS estimate would be for a point in time that was 3.5 years before the DHS survey.

### 2.3. Direct versus Indirect Estimates

Another potential source of discrepancies between estimates from DHS and other sources is that the other sources may have used indirect estimation. This is of particular relevance for comparisons with estimates from a census or survey that did not collect a birth history. Indirect estimation techniques were developed for such sources primarily during the 1950s and 1960s, before it had been demonstrated that it was possible to collect birth histories of adequate quality in most settings. Methods were developed to estimate a total fertility rate using the number of children ever born for women in successive 5-year age groups, and infant/child mortality rates using the number of children ever born and the number still alive for women in successive 5-year age groups. These methods obtain their leverage from a combination of assumptions, such as an assumption that fertility or mortality has not changed or has changed linearly in recent years, and synthetic data sources, such as model life tables or model fertility schedules (see UN Manual on Indirect Estimation (United Nations 1983)). There is a general, although not universal, consensus that direct estimates are more accurate than indirect estimates, except if the birth histories have very high levels of omission or displacement. When there is omission of births and/or deaths, reports of the numbers of children used for the indirect estimates will also be low. If there appears to be a discrepancy between a DHS direct estimate and an indirect estimate from another source, it is quite possible that the assumptions and synthetic data used in the construction of the indirect estimate were inappropriate.

### 2.4. Unadjusted versus Adjusted Estimates

As stated earlier, DHS estimates are not adjusted, but comparisons will be affected by any adjustments to non-DHS rates that are being compared. For example, WHO estimates of the MMR for countries that have conducted DHS surveys make extensive use of DHS data but they apply several adjustments, including a deduction for deaths that were pregnancy-related (in terms of timing) but not obstetric, and an increase to compensate for likely omission. Both adjustments are based on limited evidence. The model is Bayesian, and borrows information from the UN regional level. The model uses covariates, including GDP estimates from the World Bank, and it applies the DHS estimate of the proportion of deaths to women that were maternal to an estimate of the number of deaths to women that comes from the UN Population Division. The use of UN Population Division adjusted estimates of age-specific death rates for adults, rather than DHS unadjusted estimates of these rates, as well as the UNPD adjusted estimate of the General Fertility Rate for the denominator of the MMR, account for the differences between DHS and WHO estimates.

The IGME estimates of infant and child mortality rates involve a blending of many sources. The model for those estimates does not include any covariates. Detailed documentation on the adjustments made to the alternative estimates are readily available at the websites provided elsewhere in this report.

### 2.5. Full Histories versus Truncated Histories

An additional source of difference in published survey results relates to how the data are collected and processed. All PMA2020 and most MICS surveys, for example, simplify data collection by using abbreviated birth histories for specified intervals before the interview that range from 1 to 5 years and by not collecting data on multiple births. When a DHS estimate is compared with an estimate from a truncated birth history, there will be obvious differences such as fewer births per respondent and a shorter reference period. There may be other limitations as well. For example, when respondents are asked to include events in a time interval, such as the past year, it is not clear if the nominal interval of 1 year is consistently interpreted by all respondents and interviewers. However, the focus on very recent events may reduce omission. Alternate methods for imputing incomplete or unknown dates may also cause variation in published fertility rates. Chapter 4 of this report will focus on a strategy for comparisons of this sort that uses microdata files and simulation from PMA2020 and DHS data.

# 3. Strategies for Comparisons

This chapter will illustrate methods for comparing DHS estimates with one another, or comparing a DHS estimate with another estimate, making a judgement about consistency and plausibility, and taking into account several potential reasons why the estimates could differ, including sampling variation. In some situations, the alternative estimates can be judged to be consistent. In other situations, the DHS estimate—or perhaps the alternative estimate—may appear to be biased in one direction or another. If reconciliation is not possible, then the user must decide whether to blend them with some statistical procedure or to prioritize one estimate over another.

## 3.1. Data and Indicators

### *Data*

This chapter will use data from 51 DHS surveys. These are surveys for which fieldwork began in 2010 or later and for which the data had been released by the closing date for this report. If a country had conducted more than one survey in that interval, only the most recent survey was used. A list of these surveys and the reference dates of the indicators is shown in Table A.1 of the Appendix. All estimates are at the national level. The DHS estimates used in this chapter are available on STATcompiler, with the exception of some confidence intervals as described below.

Comparisons will be made with estimates from three other sources, described with acronyms: IGME (UN Inter-Agency Group for Child Mortality Estimation); UNPD (UN Population Division); and WHO (World Health Organization). These acronyms are convenient but misleading, because the three sources are part of the United Nations and are inter-related. The IGME estimates are associated with UNICEF (the UN Children's Fund), and the WHO estimates involve an inter-agency group known as the MMEIG (Maternal Mortality Estimation Inter-Agency Group). The websites are:

IGME: *https://data.unicef.org/topic/child-survival/* (UNICEF 2017)

UNPD: *https://esa.un.org/unpd/wpp/* (United Nations 2017)

WHO: http://www.who.int/reproductivehealth/publications/monitoring/maternal-mortality-2015 (WHO 2015)

All the IGME, UNPD, and WHO data used in this report are freely available on these websites.

This chapter also single out two specific DHS surveys: the survey conducted in Afghanistan in 2015 and the survey in Angola in 2015-16. The results from both surveys became available in 2017. There have been some data quality concerns for both surveys, but they are selected mainly to provide more specificity to a strategy for identifying deviations from what could have been expected.

### *Indicators*

Six indicators will be used: the Total Fertility Rate (TFR), the Infant Mortality Rate (IMR), the Under-5 Mortality Rate (U5MR), the Adult Female Mortality Probability (AFMP), the Adult Male Mortality Probability (AMMP), and the Maternal Mortality Ratio (MMR). All are synthetic (that is, they are calculated with recent data but have a cohort interpretation) and involve complex calculations, as described below.

Total Fertility Rate (TFR): The sum of the age-specific fertility rates for ages 15-49 (multiplied by five if these are 5-year rates), using data from the birth histories for the 3 years before the survey. The TFR can be interpreted as the average number of births that a woman would have if she survived from the 15th birthday to the 50th birthday and had children according to current age-specific rates. The TFR estimates used for this chapter are summarized in Table A.2.

Infant Mortality Rate (IMR): The probability that a child will die before the first birthday (1q0), multiplied by 1000. DHS uses data from the birth histories for the 5 years before the survey to calculate the IMR. See Table A.3.

Under-5 Mortality Rate (U5MR): The probability that a child will die before the fifth birthday (5q0), multiplied by 1000. DHS uses data from the birth histories for the 5 years before the survey to calculate the U5MR. See Table A.3.

Adult Female Mortality Probability (AFMP): The probability that a woman who has survived to her 15th birthday will die before the 50th birthday (35q15), multiplied by 1000. DHS uses data from the sibling histories for the 7 years before the survey to calculate the AFMP. See Table A.4.

Adult Male Mortality Probability (AMMP): The probability that a man who has survived to his 15th birthday will die before the 50th birthday (35q15), multiplied by 1000. DHS uses data from the sibling histories for the 7 years before the survey to calculate the AMMP. See Table A.4.

Maternal Mortality Ratio (MMR): The number of maternal deaths per 100,000 births. DHS obtains maternal deaths from the sibling histories to calculate a MMR and births from the birth histories to calculate a General Fertility Rate (GFR), by using the 7 years before the survey, dividing the MMR by the GFR, and multiplying by 100,000. The MMR can be interpreted roughly as the probability that a pregnancy will result in a woman's death (multiplied by 100,000). Births, rather than pregnancies, are used in the denominator because the number of pregnancies is not available. See Table A.5.

The IMR and U5MR have "rate" in their names but are actually probabilities. The AFMP and AMMP are also probabilities, not rates. As a group, these four probabilities are preferred over rates, because most rates are affected by the age distribution, while the probabilities are not. The TFR is a rate but is not affected by the age distribution. The MMR is a ratio of two rates, both of which are affected by the current age distribution of women age 15-49.

All six of these indicators are provided in STATcompiler. The upper and lower bounds of a confidence interval for the MMR is given in Stat Compiler. The estimates of all the indicators, shown in Tables A.2 – A.5 of the Appendix, and the confidence interval for the MMR, are available in STATcompiler. At this time, STATcompiler does not include confidence intervals for the other indicators, although it is likely that confidence intervals will be added in the future. Appendix B of the main reports includes confidence intervals for the TFR, IMR, and U5MR. Special computer runs for this chapter re-calculated those confidence intervals, using adjustments for sampling weights, clustering, and stratification. Confidence intervals for the AFMP and AMMP were not calculated.

The DHS estimates will be compared with IGME estimates of the IMR and U5MR; UNPD estimates of the TFR, AFMP, and AMMP; and WHO estimates of the MMR. The analysis is based on a set of 21 figures provided in the Appendix, which in turn are based on tables also provided in the Appendix that follows the figures.

### 3.2. Illustrative Comparisons

We begin with an overview of the structure of the figures. The following list specifies which figures apply to each of the six indicators:

| Indicator | Figure number |
|---|---|
| TFR | A.1, A.2 |
| IMR | A.3, A.4, A.7, A.8, A.9 |
| U5MR | A.5, A.6, A.7, A.8, A.9, A.14, A.15, A.20, A.21 |
| AFMP | A.10, A.12, A.13, A.14, A.15, A.18, A.19 |
| AMMP | A.11, A.12, A.13 |
| MMR | A.16, A.17, A.18, A.19, A.20, A.21 |

Many figures are paired with another, in such a way that the first figure focuses on the estimated levels of the DHS and UN indicator, and the second on the difference between the two estimates. Figures A.7, A.10, and A.11 are the only figures that are not part of such a pairing.

About half of the figures involve a pairing of indicators. The following list identifies such pairs:

| Indicators | Figure number |
|---|---|
| IMR and U5MR | A.7, A.8, A.9 |
| AFMP and AMMP | A.12, A.13 |
| U5MR and AFMP | A.14, A.15 |
| MMR and AFMP | A.18, A.19 |
| U5MR and MMR | A.20, A.21 |

This pairing of indicators is based on empirical correspondences. The IMR and U5MR have a logical or definitional association because they are the probabilities of dying before age 1 and the probability of dying before age 5, respectively. It is a formal requirement, therefore, that IMR<U5MR. Beyond that, the two probabilities are highly correlated empirically. Figures A.7, A.8, and A.9 describe this pairing. The adult probabilities of dying between age 15 and 50 (AFMP and AMMP) will be associated because many causes of death for women and men are the same. This pairing is examined in Figures A.12-A.13.

There are also definitional constraints on the relationship between the MMR and the AFMP, because a maternal cause of death is one of many possible causes of death. This pairing is described in figures A.18-A.19. The MMR and AFMP are on different scales, and there is not a simple inequality, although there is a relationship.

Two pairings of indicators are based on the commonality of many influences on the mortality of children and women. We thus examine the pairing of the U5MR with the AFMP, in Figures A.14-A.15, and with the MMR, in Figures A.20-A.21.

Some relationships are clearer on a log scale than on the original scales of rates and probabilities. Figures A.8 and A.9 use the same data as Figure A.7, but on a log scale. Figure A.9 is a differencing of Figure A.8. Logarithms have a base of 10 (10 and 100 on the original scalar, for example, correspond with 1 and 2 on the log scale).

We will now demonstrate how these figures can be interpreted and used to make decisions about plausibility, potential data quality issues, or a need for further analysis. For each figure, we will make some general comments and will comment specifically on the Afghanistan and Angola surveys. Since there is only one survey per country, it will be sufficient to identify each survey with just the name of the country, although it must be understood that the reference is to a specific survey. We will show confidence intervals for many of the DHS estimates but will not conduct statistical tests. However, if the other estimate is outside the DHS 95% confidence interval, and we treat the other estimate as a fixed point, it will be safe to say that there is a significant difference with $p<.05$. Numerical values can be made more precise by referring to the tables in the Appendix.

### *Total Fertility Rate*

Figure A.1 compares the TFR estimate with the TFR estimate from the UN Population Division (identified as the “UN” estimate). The UN estimate is for July 1 of the calendar year that is closest to the reference date (the midpoint of the reference period) for the DHS estimate. The DHS estimate is represented by a blue interval—the 95% confidence interval. The DHS point estimate would be at approximately the midpoint of the confidence interval. The countries are sorted by the DHS point estimate. The UN point estimate is shown with a red dot.

The DHS and UN estimates are not independent. Some of the UN trajectories for the TFR were updated with the most recent survey. The UN estimates come from the 2017 revision of World Population Prospects, which was issued in June 2017. For every country, the UN Population Division does a careful analysis of the plausibility of new estimates. When a country trajectory is revised on the basis of new data, the new estimate has only a limited effect.

For most countries, the UN estimate is within the DHS confidence interval. It is easy to identify the countries in which the DHS interval is completely below or completely above the UN estimate. The numbers of these two types of displacement are approximately equal. The agreement appears to be strong, in general and specifically for both Afghanistan and Angola.

Figure A.2 maintains the same ordering of countries, on the basis of the DHS estimate of the TFR, but provides green bars that show the DHS estimate minus the UN estimate on an exaggerated scale. For all countries, the DHS and UN estimates are within half a child of each other. The largest differences are for the Kyrgyz Republic, Haiti, and Mozambique for which the DHS estimate was higher than the UN estimate, and Ethiopia and Malawi, for which the DHS estimate was lower than the UN estimate. With the Afghanistan and Angola surveys, the agreement is good for both countries.

### *Infant Mortality Rate*

Figure A.3 is a similar comparison that uses the IMR. The non-DHS source is IGME. The majority of IGME estimates are within the DHS confidence interval, although there are a number of surveys for which the IGME estimate is completely above the DHS confidence interval. For most countries where the DHS estimate of the IMR is above 50, the IGME estimate is significantly higher. The greatest weakness of estimates of infant and child mortality from a birth history is the potential omission of children who died. These omissions have a minor effect on fertility rates but a major effect on mortality rates. Omissions are almost certainly greatest in the countries with the highest prevalence of early deaths.

In Figure A.3, the IGME estimate for Angola is higher than the IGME estimate for any other country, with the exception of Sierra Leone. Extreme values of any indicator, from any source, should be treated with caution. It is possible that the high deviation for Angola is partially due to a potential IGME over-estimate of the true IMR.

The IMR estimates for Angola are further apart than for any other country. The IGME estimate is more than twice the DHS estimate. The deviation for Afghanistan is not as great as that for Angola, but is among the six largest deviations.

Figure A.4 gives the "DHS minus IGME" differences on their own scale, with more clarity than in Figure A.3. As in Figure A.3, the countries are sequenced by the DHS estimate of the IMR. Six countries have a deviation of more than 20 points, including Afghanistan and especially Angola. Apart from the possibility that the IGME estimates for these two countries, particularly Angola, are too high, it appears that the DHS estimate of the IMR for Angola is much too low, by perhaps 50 points, and the estimate for Afghanistan is somewhat low, by perhaps 20 points.

### *Under-5 Mortality Rate*

The U5MR is highly correlated with the IMR. There is a close correspondence between Figures A.5 and A.3 and between Figures A.6 and A.4, although a comparison of Figures A.5 and A.3 shows a better correspondence between DHS and IGME for the U5MR than for the IMR. Only six of the 51 countries have conspicuous deviations. In addition to Afghanistan and, especially, Angola, these are Comoros, Gambia, Benin, and Mali. For these six countries, the IGME estimate is at least 28 points higher than the DHS estimate. For Angola, there is a 100 point difference.

According to the IGME estimates, the true U5MR for Angola is the highest of 51 countries. Chad and Sierra Leone are the only countries included here for which the IGME estimate is higher than 140. If correct, the level of omission in the Angola survey was enormous, with more than half of under-5 deaths omitted.

### *Comparing the Infant and Under-5 Mortality Rates*

Figure A.7 is a scatterplot with the IMR on the horizontal axis, the U5MR on the vertical axis, and countries (or surveys) represented by dots. Blue dots represent the DHS estimates and red dots the IGME estimates. Figure A.8 matches exactly with Figure A.7, except that in Figure A.8, both rates have been logged to reduce the compression of the lower rates in Figure A.7. The figures include regression lines, which were calculated separately for the DHS and IGME estimates.

As stated above, the IMR and U5MR are highly correlated. The correlation squared on the original scales of the rates, in Figure A.7, is 0.88 for the DHS estimates and 0.96 for the IGME estimates. It is even higher on the log scales. In Figure A.8, the correlation squared is 0.94 for the DHS estimates and 0.98 for the IGME estimates. In both figures, the IGME correlation is somewhat greater than the DHS correlation, probably because the IGME adjustments are designed, in part, to improve the correspondence between the IMR and U5MR.

DHS surveys that deviate substantially from either the blue line or the red line, in either of Figures A.7 or A.8, may merit further investigation. For example, in Figure A.7, a blue dot stands out, with the greatest deviation above the blue regression line for a survey with an IMR that is substantially lower than would be expected, given its U5MR. This point represents the Niger 2012 survey. A possible explanation for such a deviation is that a high proportion of infant deaths were misclassified by being reported for age 1 rather than age 0. According to Table A.3 in the Appendix, the DHS estimates of the IMR and U5MR were 51 and 127, respectively. The IGME estimates are 69 and 132 respectively, with the main difference a re-allocation of deaths from age 1-4 into age 0.

Another blue dot represents the greatest deviation below the blue line in Figure A.7, with a U5MR that is substantially lower than expected, given its IMR. This dot represents Pakistan, which has DHS estimates for the IMR and U5MR of 74 and 89, respectively. The IGME estimates are 74 and 92, which does not change the IMR estimate but moves the point upward toward the red line, which is below the blue line.

Figures A.7 and A.8 also include labels for the blue dots for the DHS estimates for Afghanistan and Angola and red dots for the IGME estimates for those two countries. The blue dots for those two countries do not conspicuously deviate from the regression lines. Thus, the combination of the IMR and U5MR is plausible.

However, the red dots for those two countries are substantially above the corresponding blue dots, on both scales, especially for Angola.

In Figures A.7 and A.8, it would be helpful to be able to pair the blue and red dots for specific countries—that is, to know how the difference between the DHS and IGME estimates of the U5MR relates to the difference between the DHS and IGME estimates of the IMR for the same country. Figure A.9 describes those differences, working from Figure A.8, which is on a log-log scale. In Figure A.9, the horizontal axis is the difference between the values of corresponding blue and red dots on the horizontal axis of Figure A.8. The vertical axis is the difference between the values of corresponding blue and red dots on the vertical axis of Figure A.8.

***Female and Male Probabilities of Dying Between 15 and 50***

The remaining DHS estimates are limited to 33 surveys/countries that included sibling histories and produced adult mortality estimates. They are compared with other estimates of adult mortality for women and men in those countries at approximately the midpoint of the reference period produced by the UN Population Division, referred to as the UN. Figure A.10 compares the UN and DHS estimates for women with a scatterplot. There is high correlation. The square of the correlation is 0.68, and the regression line is shown. The figure also includes a 45 degree line for equality of the DHS and UN estimates. Figure A.10 identifies countries for which there is a high deviation from either the regression line or the line of equality. Most points are above the line, which indicates that the UN estimate is higher than the DHS estimate, because of a UN adjustment for omission, in the sibling histories, of sisters who died between the age 15 and 50. There are a number of upward adjustments of .10 (10 percentage points) or more, which will not be discussed. There is a surprising deviation for Zambia, for which the UN estimate is negative by about 10 percentage points. A downward adjustment, especially by this magnitude, is unusual. Afghanistan and Angola have very similar estimates from both DHS and the UN, with the points for both almost exactly on the line of equality.

Figure A.11 is analogous to Figure A.10 but for men. The DHS estimates are based on the brothers in the sibling histories. The interpretations are almost exactly the same as for women in Figure A.10. Most of the deviations from the line of equality are in the direction of a higher estimate from the UN than from DHS. Zambia stands out with a large deviation in the opposite direction, while the dots for Afghanistan and Angola are almost exactly on the line of equality of the DHS and UN estimates.

Figures A.12 and A.13 compare the probabilities of dying between age 15 and 50 for women and men, as estimated by DHS (blue dots and line) or the UN (red dots and line). In this figure, a diagonal line corresponds with equal probabilities for women and men. The DHS and UN regression lines are very close to each other. The correlation squared is very high, at 0.86 for DHS and 0.90 for the UN. Both lines have a slope close to 1 and are just slightly above the line of equality. That is, female mortality is higher than male mortality about as often as it is lower, but is a little more likely to be lower. There is a great deal of variation in these probabilities, although the rates for women and men are usually quite close.

With Afghanistan and Angola, the figure shows a relatively large deviation for the Afghanistan DHS, which is below all lines. The male rate is low, given the female rate or, conversely, the female rate is high, given the male rate. Since under-reporting of deaths is much more likely than over-reporting, there appears to be omission of male deaths. The corresponding red dots, for the UN estimates for Afghanistan and Angola, are closer to all lines.

The differences between the DHS and UN estimates are more easily seen with Figure A.13, which plots the differences for women on the horizontal axis and for men on the vertical axis. Afghanistan is represented by a point at approximately (-.05,-.01); the DHS estimate is below the UN estimate by .05 for women and below by .01 for men. That is, the DHS estimates are low (relative to the UN estimates) but only by a negligible amount (.01) for women and a larger amount (.05) for men. In Figure A.13, the point for Angola is very close to the origin (0,0), which implies that it almost perfectly matches the UN estimate. There are many other surveys for which the DHS and UN estimates of adult mortality differ by much more than those for Afghanistan and Angola.

### *Comparing Adult Female Mortality and Under-5 Mortality*

Figure A.14 is another scatterplot with blue dots for DHS estimates and red dots for UN estimates. The horizontal axis is the U5MR and the vertical axis is the AFMP. Two regression lines are shown. The lines are very close to each other but the fit is relatively poor, compared with the regression lines in earlier figures. These two estimated probabilities are not strongly associated. There are several outliers from the general pattern, but these would be identifiable in the previous univariate figures for the U5MR and the AFMP. Figure A.15 plots the differences between the DHS and UN estimates, with the U5MR difference on the vertical axis and the AFMP difference on the horizontal axis. Afghanistan and Angola stand out because the DHS estimates of the U5MR are much lower than the UN estimates. The two points are very close to the vertical axis because the DHS and UN estimates of the AFMP agree closely. Other outliers in this figure would have been identified earlier.

### *Maternal Mortality Ratio*

Figure A.16 compares 31 DHS estimates of the MMR, including confidence intervals, with the WHO point estimates. WHO provides uncertainty intervals, but these are model-based and cannot be interpreted as confidence intervals. The figure includes the 31 DHS surveys with sibling histories. The dates of the WHO estimates are approximately at the midpoints of the reference intervals for the DHS seven-year estimates, and not at the dates of the surveys.

Again, the WHO estimates are not independent of the DHS estimates, because DHS data are used in the construction of the WHO estimates. The WHO estimates have had multiple adjustments that are not applied to the DHS estimates.

For 16 surveys, the WHO estimate is within the confidence interval for the DHS estimate. For 10 surveys, the WHO estimate is above the DHS confidence interval and for 5 it is below. The differences between the two estimates, expressed as the DHS estimate minus the WHO estimate, are shown in Figure A.17. In this figure, a majority of the differences are within the DHS confidence interval.

The Angola survey has one of the most serious under-estimates, relative to the WHO estimate. There are eight surveys in which the DHS point estimate is at least 200 points below the WHO point estimate. The negative deviation for Angola is the same as for Nigeria, and is exceeded only by Gambia and Burundi.

It is less common for the DHS estimate to be greater than the WHO estimate. A positive deviation greater than 200 points is observed for only three countries—Zimbabwe, Lesotho, and, largest of all, Afghanistan. The deviation for Afghanistan is the greatest, by far, among the 31 countries.

Figures A.18 and A.19 compare the DHS and UN/WHO estimates of adult female mortality and maternal mortality. This kind of comparison focuses on the balance of adult female deaths that are classified as maternal or not maternal. Figure A.18 is a scatterplot in which the blue dots represent the combination of the two for DHS and the red dots the combination for UN/WHO estimates. The horizontal axis is the AFMP for women, which is a probability, and the vertical axis is the MMR, which is a complex ratio. The horizontal and vertical scales are quite different, but are used because of the measures that are readily available from general websites. The correlations between the two indicators are moderately high for both the DHS estimates and the UN/WHO estimates, and the two fitted lines are almost perfectly aligned with each other, although the blue and red scatterplots have a number of differences.

The interpretation of Figure A.19, which plots the differences between the two sets of estimates, with one dot for each county, is easier to interpret. There are several outliers on this scatterplot. Points that are relatively far to the left or the right identify countries with the most substantial DHS versus UN differences in the AFMP estimates. Points that are relatively far upwards or downwards identify countries that have substantial DHS versus WHO differences in MMR estimates. Figure 19 identifies countries with all four possible combinations of deviations.

### *Comparing Adult Female Mortality and Maternal Mortality*

With Afghanistan, the AFMP estimates from DHS and the UN are very close, although the MMR estimate is much higher for DHS than WHO. This combination of differences implies that a substantial number of deaths to women were misclassified as maternal. This interpretation is much more easily reached with Figure A.19 than with the separate figures for the AFMP and the MMR.

Angola is somewhat low for both the AFMP and the MMR estimates, but more for the MMR estimate than the AFMP estimate. If these differences are interpreted as errors, DHS missed some deaths for women, and did not correctly classify some of the women's deaths that were actually maternal.

### *Comparing Under-5 Mortality and Maternal Mortality*

A final comparison between child mortality and maternal mortality is shown in Figures A.20 and A.21. Figure A.20 is another overlay of two scatterplots, one with blue dots for DHS data and one with red dots for IGME/WHO data. In this figure, the fitted lines are quite different. The fitted line based on IGME/WHO data has a much better fit than the DHS line. Figure A.21 plots the differences in the U5MR dimension (the vertical axis) against the differences in the MMR dimension (the horizontal axis). Most of the points are in a relatively narrow range, although about a third could be described as outliers that warrant further analysis. The two most serious outliers, by far, are Afghanistan and Angola. Both of these countries have low estimates of the U5MR, relative to the IGME estimates. Angola has the most extreme deviation. Angola has a lower MMR than expected, by comparison with the WHO estimate, while Afghanistan is very high.

Figures A.1-A.21 show some redundancies. Some are more helpful than others, but together they provide a strategy for comparing DHS estimates with the various IGME, UN Population Division, and WHO estimates, calibrated with reference dates.

# 4. Comparability of Fertility Estimates from DHS and non-DHS Surveys: The Case of PMA2020 Data

In countries that lack comprehensive vital registration systems, fertility rates are often computed from nationwide surveys that ask women of reproductive age about their birth history. DHS surveys are the largest such worldwide source of survey data on fertility in low- and middle-income countries. In recent years, a number of independent surveys have also produced nationwide fertility estimates. When rates differ by survey source, it may be difficult to know which is more accurate or which should be used for trend analysis. This chapter explores variation in extant and simulated fertility rate estimates between DHS and non-DHS surveys, and uses Performance Monitoring and Accountability 2020 (PMA2020) as a reference point for estimates and methodology.

As with any survey indicator, a number of factors affect fertility estimates. First and foremost, there is natural variation in sampling and response rates that would be expected to produce slightly different estimates even if a survey was conducted in the exact same way at the exact same time. Added to this natural variation are the facts that DHS and non-DHS surveys are rarely conducted in the exact same months, and tend to employ slightly different sampling strategies. Non-DHS surveys may phrase questions about birth history in a slightly different way, gather only a truncated birth history, omit information about multiple births (twins), or follow different imputation procedures when dates of children's birth are unknown. Even if the data gathered by two surveys are similar, the methodology to compute fertility rates may be slightly different. DHS surveys compute fertility based on a full birth history and a 3-year reference period, while other surveys may ask for a truncated birth history and/or employ a shorter reference period. Two examples of this difference are MICS surveys, which vary in the level of birth history they gather and in their reference period for computing rates, and the PMA2020 surveys, which gather the two most recent births and use a 2-year period of exposure in computing fertility. In this analysis, we explore the real and simulated differences between the DHS and PMA2020 fertility estimates.

## 4.1. Data and Methods

The data for this chapter come from 256 DHS surveys fielded from 1985 until the present, plus five recent nationwide PMA2020 surveys whose timeframe most closely corresponds with a published DHS survey. The PMA2020 surveys are multistage cluster surveys of households, women, and service delivery points that employ a standard methodology and have been conducted in ten countries to date (Zimmerman et al. 2017). In comparing DHS and PMA2020 data and methods, we focus on two key fertility measures: the adolescent fertility rate and the total fertility rate (TFR). The adolescent fertility rate is measured by births per 1,000 women age 15 to 19, and corresponds to the number of average annual live births per woman age 15 to 19 during a specified reference period before the date of the survey. The TFR is a synthetic measure of the average number of children a woman would have between age 15 and 49 if she gave birth at prevailing 5-year age-specific rates.

We begin by comparing PMA2020 adolescent fertility rates and total fertility rates to corresponding DHS estimates for five countries. The PMA2020 data on Ghana were excluded in response to concerns about the plausibility of their fertility estimates.[5] DHS fertility rates were computed with the tfr2 package in Stata (Schoumaker 2013) that replicates DHS methodology, including the standard 3-year reference period. We refer to these standard DHS estimates as a *"3-year n-birth estimate"* because they use a window of 3 years prior to interview and allow as many births as were reported in those 3 years, including multiple births.

[5] Authors' discussion with PMA2020 staff, August 2017.

PMA2020 collects up to the two most recent births, does not record whether the births are multiples or not, and uses a post-estimate inflation factor to adjust fertility rates for multiple births. To compute PMA2020 fertility rates, we used the tfr2 package and then added an age-specific adjustment factor for multiple births derived from DHS data on children born in the past 5 years. Specifically, we examined all single or the first of multiple live births reported to DHS in the 5 years before the survey, and computed the percentage of births that were the first of multiple births by 5-year age group. This method, as employed by PMA2020, does not consider whether a multiple birth was a triplet or more. Multiple births greater than two are rare, and are not expected to influence the number greatly. We used the age-specific percentage of live births to women age 15 to 49 in the past 5 years that were multiple births to inflate the age-specific fertility rates. These age-specific adjusted rates were then used to compute a new TFR and to adjust the standard errors for all rates. Fertility rates computed by this method are referred to as *"2-year 2-birth adjusted rates"* because they consider up to two births and a reference period of the past 2 years, and also adjust for multiple births. All results in this chapter were computed with survey-specific weights provided with the dataset.

Next, because the extant PMA2020 data offer only five points of comparison, we compare DHS results from the standard "3-year n-birth" method to the simulated "2-year 2-birth adjusted fertility rate" estimates using all 256 available standard, continuous, and interim DHS datasets from 1985 onward. In order to make the DHS data comparable to PMA2020, we dropped any birth that was more than the first of a multiple set, kept the two most recent births, and used a 2-year reference period for fertility rates. We added an age-specific adjustment factor for multiple births as described above and proceeded with DHS as with PMA2020 data to adjust the fertility rates and standard errors. How do the rates and standard errors using a 2-year 2-birth adjusted method compare to the traditional 3-year n-birth TFR published by DHS? Do these estimates always overlap?

Finally, it is important to consider that non-DHS surveys use a different methodology for computing fertility rates, and may use a different sampling strategy. In particular, estimates of fertility from non-DHS surveys may be based on a smaller number of clusters. In the third section of this chapter we simulate 2-year 2-birth adjusted fertility rates with 2016 Ethiopia DHS data 1,000 times using a sub-sampling strategy that follows from a recent PMA2020 survey in Ethiopia (Round 4). For the sampling strategy, we compared PMA2020 cluster selection to the DHS cluster selection by strata—urban or rural residence within region with some regions grouped—and randomly sampled the number of PMA2020 clusters from among the DHS clusters in the same strata. In cases where PMA2020 had selected a greater number of clusters in the urban area of a region, all corresponding DHS clusters were selected and a compensatory number of clusters in the corresponding rural area were selected. Data were re-weighted to equal the original number of DHS respondents in the stratum. For each subsample, we computed 2-year 2-birth adjusted fertility rates and compared the resulting point estimate and standard error to the nationwide sample. One thousand sample simulations were run. We examine how these estimates of rates and standard errors compare to the 3-year n-birth estimate provided by the full 2016 Ethiopia DHS sample, and discuss the limitations of this approach.

## 4.2. Comparison of PMA2020 and DHS Fertility Estimates

We examined data from five national surveys conducted by PMA2020 in Ethiopia, Indonesia, Kenya, Nigeria, and Uganda and compared their fertility results to DHS surveys that were conducted within 3 years of the PMA2020 survey. The resulting estimates, standard errors, and sample sizes are shown in Table 4.1. As the table indicates, the TFR estimates from PMA2020 surveys tend to be lower than those from DHS by 5% (Uganda) to 22% (Nigeria), while adolescent fertility rates range from 17% lower (Nigeria) to 4% higher (Kenya) than corresponding DHS estimates.

**Table 4.1. Fertility rate results from five DHS and PMA2020 surveys**

| | DHS | | | | | | PMA2020 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Country | Year | TFR | SE of TFR | Adolescent FR | SE of Adolescent FR | Sample size | Year and round number | TFR | SE of TFR | Adolescent FR | SE of Adolescent FR | Sample size |
| Ethiopia | 2016 | 4.6 | 0.06 | 79.5 | 2.8 | 15,683 | 2016 (Round 4) | 4.2 | 0.10 | 77.3 | 4.8 | 7,481 |
| Indonesia | 2012 | 2.6 | 0.03 | 48.5 | 1.6 | 45,607 | 2015 (Round 1) | 2.3 | 0.06 | 48.0 | 4.1 | 10,455 |
| Kenya | 2014 | 3.9 | 0.04 | 96.3 | 2.4 | 31,079 | 2014 (Round 2) | 3.4 | 0.11 | 99.7 | 7.6 | 4,329 |
| Nigeria | 2013 | 5.5 | 0.04 | 121.7 | 2.3 | 38,948 | 2016 (Round 3) | 4.3 | 0.08 | 101.5 | 4.8 | 11,054 |
| Uganda | 2011 | 6.2 | 0.09 | 134.5 | 4.9 | 8,674 | 2014 (Round 1) | 5.9 | 0.17 | 139.9 | 9.1 | 3,716 |

Figure 4.1 shows a chart of comparative TFRs and their confidence intervals from PMA2020 and corresponding DHS surveys. Although—with the exception of Nigeria—the TFRs are generally close, their confidence intervals do not overlap in any study country. Adolescent fertility rates and their confidence intervals are shown in Figure 4.2. Confidence intervals for the corresponding estimates overlap in every country except Nigeria.

**Figure 4.1. TFR comparison between DHS and PMA2020**

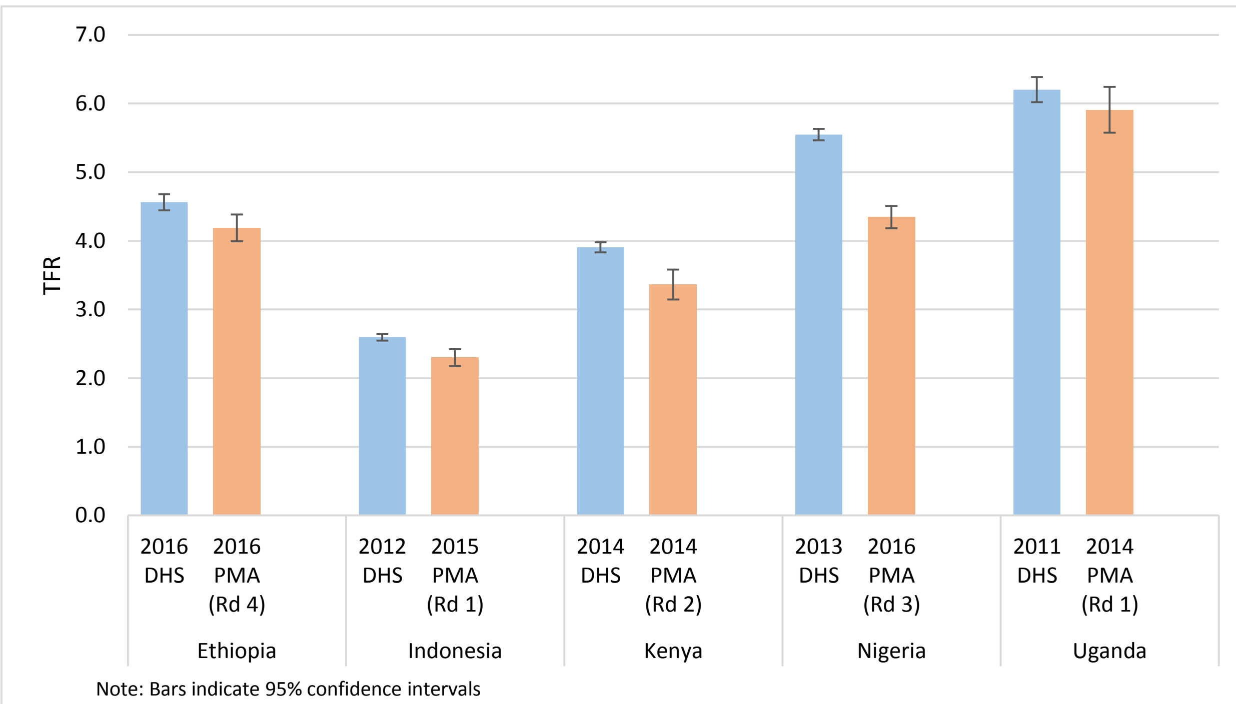
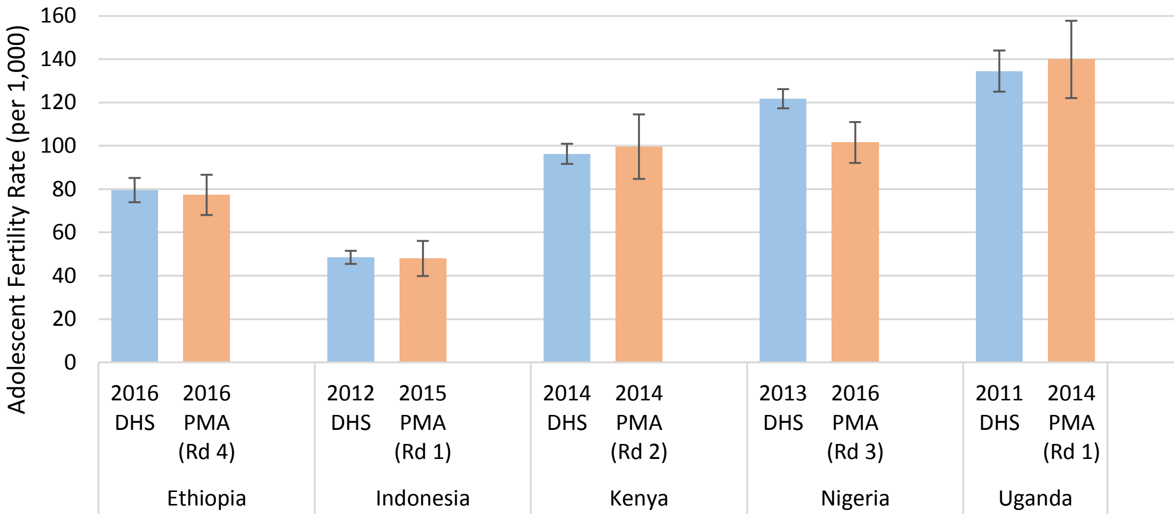


**Figure 4.2. Adolescent fertility rate comparison between DHS and PMA2020**




Note: Bars indicate 95% confidence intervals

## 4.3. A Comparison of 2-year 2-birth Adjusted Fertility Rates with 3-year n-birth Rates using DHS Data

As discussed in the introduction, differences in survey fertility results may be driven by a number of factors such as questionnaire design and sampling strategy that are quite separate from questionnaire and imputation effects. Retrospective fertility rates produced by surveys are subject to errors of omission and displacement (Pullum and Becker 2014; Schoumaker 2014) that may vary by survey. Imputation of births that appear to be heaped on a particular month may affect fertility rates by a few percentage points (Choi et al. 2017). To independently assess the difference in fertility rates expected by a 2-year 2-birth adjusted method versus a 3-year n-birth method, we used all 256 available standard, continuous, and interim DHS datasets to calculate both estimates on the same data. Employing a different methodology with the same data enables us to assess the effect of computational and collection differences apart from fieldwork factors, imputation, or sampling design. A comparison of TFRs and adolescent fertility rates computed with each method for the same survey are shown in Figure 4.3. The results show a 99.8% correlation in TFRs and a 99.5% correlation in adolescent fertility rates. The differences in adolescent fertility rates appear to be larger when the rates are higher.

To gauge the specific differences between estimates, we produced histograms of the results. Figure 4.4 shows a distribution of the percentage difference in TFR estimates between the 2-year 2-birth adjusted method and the 3-year n-birth method on the exact same data and a distribution of the percentage difference in standard errors between the two estimates. The results show a fairly close match: TFR estimates from the 2-year 2-birth adjusted method range from -7% to +7% different from the 3-year n-birth estimates with a mean difference of less than one percentage point. As expected, the method that uses a truncated birth history produces higher standard errors that range from 17% to 27% higher than their 3-year n-birth counterparts, with an average difference of 22%.

**Figure 4.3. Comparison of 3-year n-birth and 2-year 2-birth adjusted TFR and Adolescent Fertility Rate, DHS surveys 1985-present**

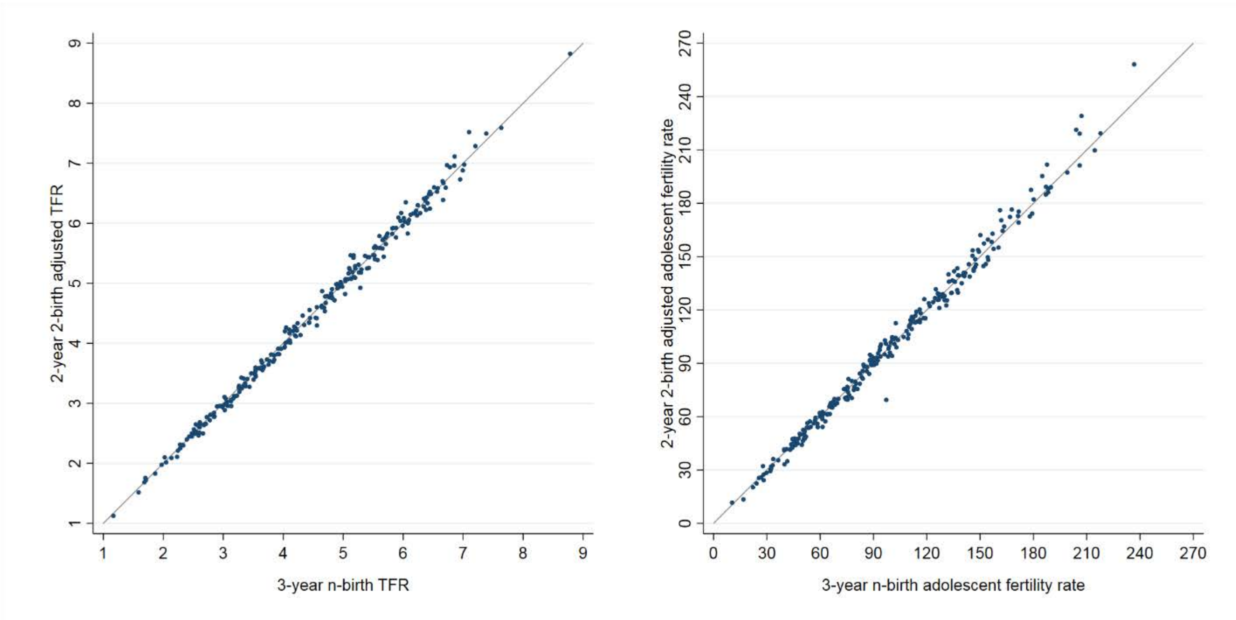
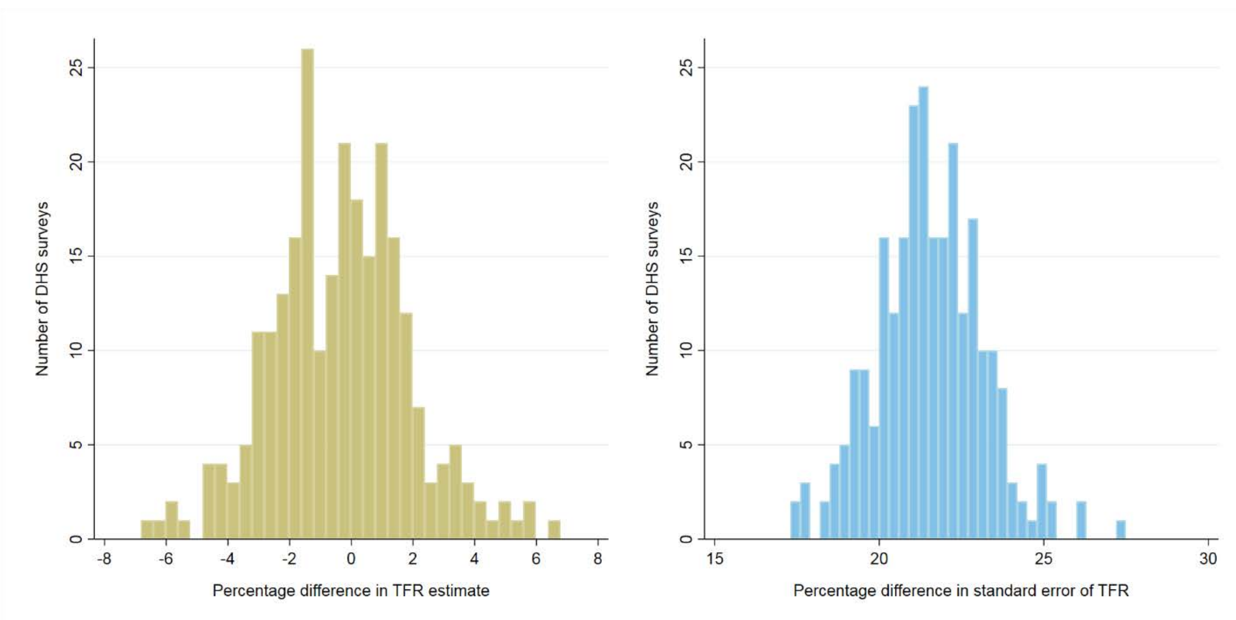



**Figure 4.4. Distribution of difference between the 2-year 2-birth adjusted TFR and the 3-year n-birth TFR estimate and standard error of TFR estimate, DHS surveys 1985-present**




The corresponding results for the difference in adolescent fertility rates and standard errors of those rates between the two computation methods are shown in Figure 4.5. Here we see more variability in rates, from a -29% to a +16% difference in rate estimates and a mean also near zero (less than one percentage point). The difference in standard errors of these estimates ranges from 1% higher to 35% higher, with an average difference in standard errors of 22%.

**Figure 4.5. Distribution of difference between the 2-year 2-birth adjusted adolescent fertility rate and the 3-year n-birth adolescent fertility rate estimate and standard error of estimate, DHS surveys 1985-present**

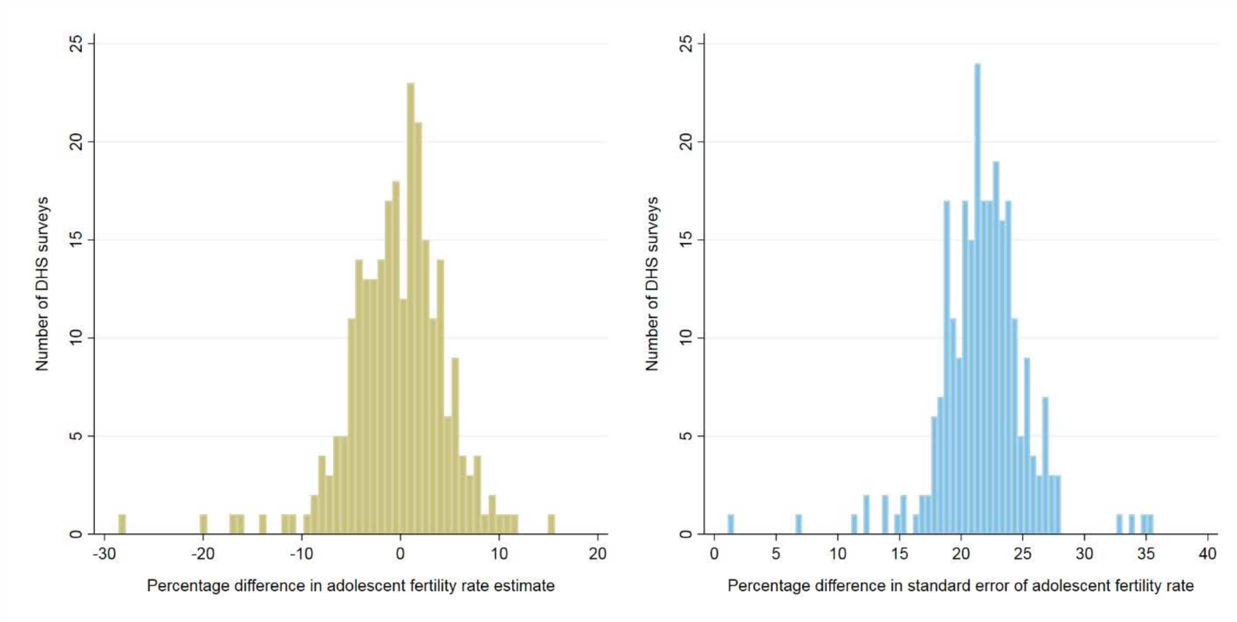
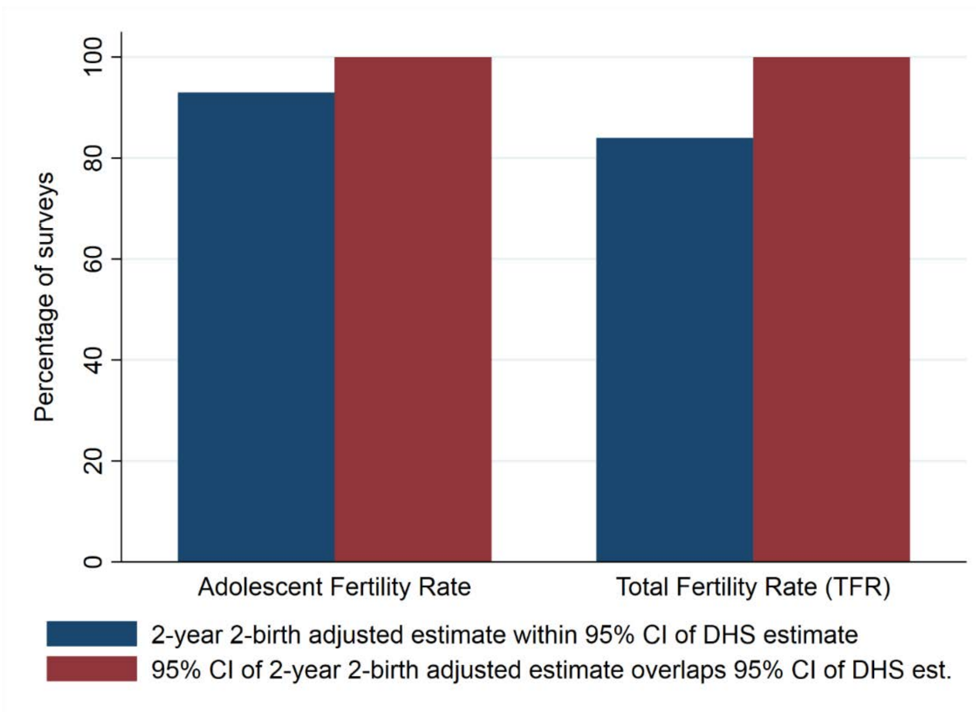



**Figure 4.6. Percentage of 2-year 2-birth fertility estimates that are within confidence bounds of 3-year n-birth estimates, DHS surveys 1985-present**




We also assessed the percentage of 2-year 2-birth adjusted estimates that fall within the confidence intervals of the DHS estimates and the percentage of the confidence intervals between the two estimates that overlap. While overlapping confidence intervals may still be statistically significantly different, this comparison provides a reasonable gauge the robustness of these alternate estimates. The results of the comparison are shown in Figure 4.6. In all, 93% of the adolescent fertility rate estimates and 84% of the TFR estimates using the 2-year 2-birth adjusted method fell within the confidence interval of the original DHS 3-year n-birth estimate. All alternate rates produced confidence intervals that overlapped with DHS estimates.

## 4.4. Simulation of 2-year 2-birth Adjusted Fertility Rates in Ethiopia under Different Sampling Conditions

In addition to alternate methods of estimating fertility between DHS and non-DHS surveys, the surveys may also have different sampling strategies. In particular, surveys that rely on a smaller number of clusters and respondents may produce different estimates of rates, particularly when a slightly different methodology is used to compute rates. In this section, we randomly simulate subsamples of the 2016 Ethiopia DHS that are more comparable to the number of clusters sampled in the comparable PMA2020 survey to explore how the use of a subsample of clusters combined with a different methodology might affect the rate estimates.

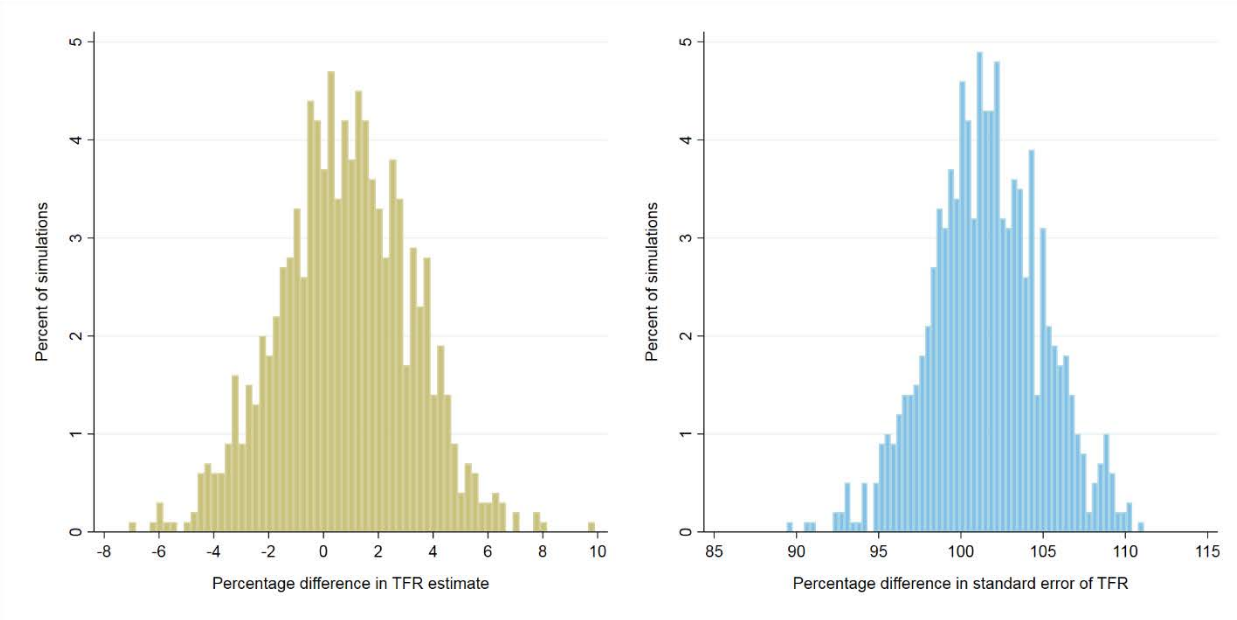
In order to produce subsamples of DHS data, we compared the number of clusters sampled by the Ethiopia 2016 DHS survey and the corresponding PMA2020 survey and chose the same stratified cluster composition as PMA2020. In the 2016 Ethiopia survey, DHS sampled 643 clusters, while the corresponding PMA2020 survey (Ethiopia Round 4) sampled 221 clusters. As shown in Table 4.2, in four regions, the Ethiopia Round 4 PMA2020 survey sampled more clusters in the urban area of the region than DHS; in these regions, a compensatory number of rural clusters were drawn. The resulting selection parameters are shown in the table; 221 DHS clusters were randomly selected from among the strata as indicated.

**Table 4.2. Distribution of clusters by strata in Ethiopia DHS 2016, Ethiopia PMA2020 2016 (Round 4), and DHS subsample simulation**

| DHS Cluster Sample - Ethiopia 2016 | | | | PMA2020 Cluster Sample - Ethiopia Round 4 | | | | Resulting Simulated Cluster Sample | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Urban | Rural | Total | | Urban | Rural | Total | | Urban | Rural | Total |
| Addis Ababa | 56 | - | 56 | Addis Ababa | 22 | - | 22 | Addis Ababa | 22 | - | 22 |
| Amhara | 11 | 60 | 71 | Amhara | 14 | 26 | 40 | Amhara | 11 | 29 | 40 |
| Oromia | 10 | 64 | 74 | Oromia | 18 | 34 | 52 | Oromia | 10 | 42 | 52 |
| SNNPR | 8 | 63 | 71 | SNNPR | 28 | 19 | 47 | SNNPR | 8 | 39 | 47 |
| Tigray | 15 | 48 | 63 | Tigray | 21 | 13 | 34 | Tigray | 15 | 19 | 34 |
| All others | 102 | 206 | 308 | All others | 6 | 20 | 26 | All others | 6 | 20 | 26 |
| | 202 | 441 | 643 | Total | 109 | 112 | 221 | | 72 | 149 | 221 |

We randomly simulated this subsample of DHS clusters 1,000 times, used a 2-year 2-birth adjusted method to compute the resulting TFR and adolescent fertility rate, and then compared the results to the overall fertility estimates from the standard 3-year n-birth full sample. The distribution of percentage differences in rates and standard errors of TFR is shown in Figure 4.5. There is fairly good consistency between rates: the 2-year 2-birth adjusted estimates of TFR using a subsample of DHS data range from a -7% to a +10% difference and average less than one percentage point difference with 3-year n-birth estimates using the full sample. However, the standard errors shown in Figure 4.7 tend to be about twice as high as those from DHS data, and range from 89% to 111% higher, with an average of 102% higher. These results compare favorably to those shown using actual PMA2020 data in Table 4.1, where the TFR was 8% lower than in the 2016 Ethiopia DHS, but appear to show a higher inflation of standard errors than in actual PMA2020 data, where the standard error of the TFR was only 65% higher.

**Figure 4.7. Distribution of difference between the 3-year n-birth TFR from Ethiopia DHS 2016 and a 2-year 2-birth adjusted TFR computed using 1,000 simulated cluster samples similar to the 2016 Ethiopia PMA2020, estimate and standard error of estimate**




Using the same 1,000 simulation subsamples, we computed the adolescent fertility rate using the 2-year 2-birth adjusted method and compared it to the standard 3-year n-birth adolescent fertility rate from the entire DHS sample. The results are shown in Figure 4.8. The percentage difference in adolescent fertility rate estimates ranges from -23% to +12% with an average of -4%. These results are consistent with the finding from the actual Round 4 PMA2020 Ethiopia survey, in which the adolescent fertility rate was 3% lower than in DHS. As with the TFR, the standard errors for adolescent fertility tend to be about twice as high in the subsample as in the corresponding DHS full sample. The errors average 97% higher, with a range of 76% to 115% higher. This estimated range is higher than the actual percentage difference in standard errors of adolescent fertility rates between PMA2020 and DHS data for Ethiopia, which were only 68% higher in the PMA2020 survey.

**Figure 4.8. Distribution of difference between the 3-year n-birth adolescent fertility rate from Ethiopia DHS 2016 and a 2-year 2-birth adjusted adolescent fertility rate computed using 1,000 simulated cluster samples similar to the 2016 Ethiopia PMA2020, estimate and standard error of estimate**

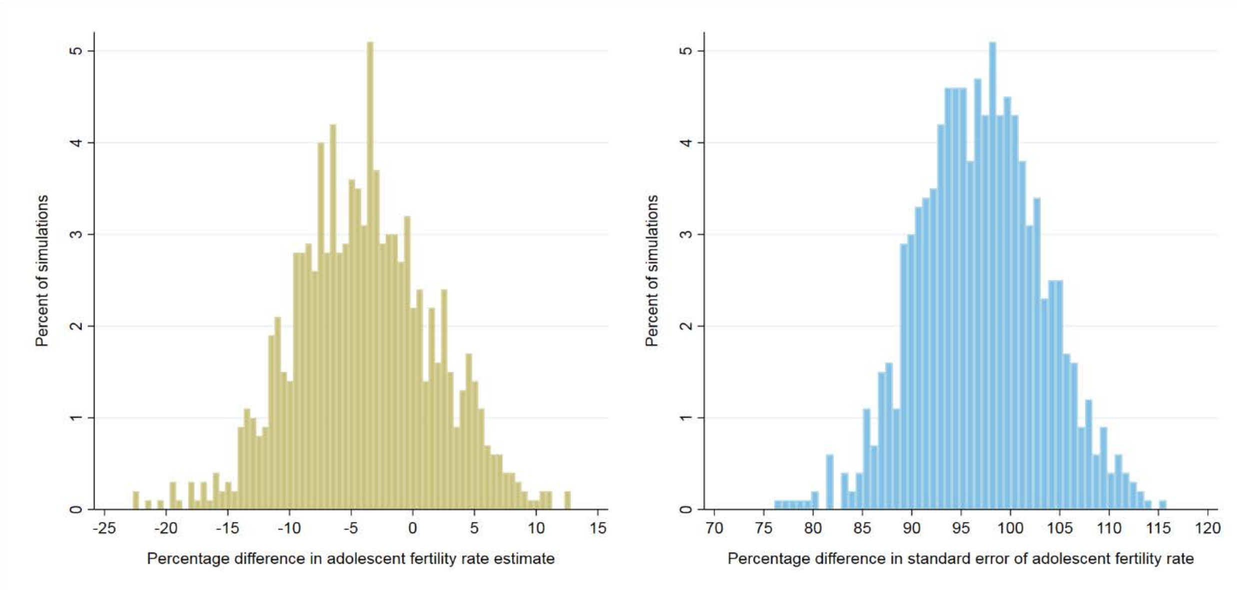
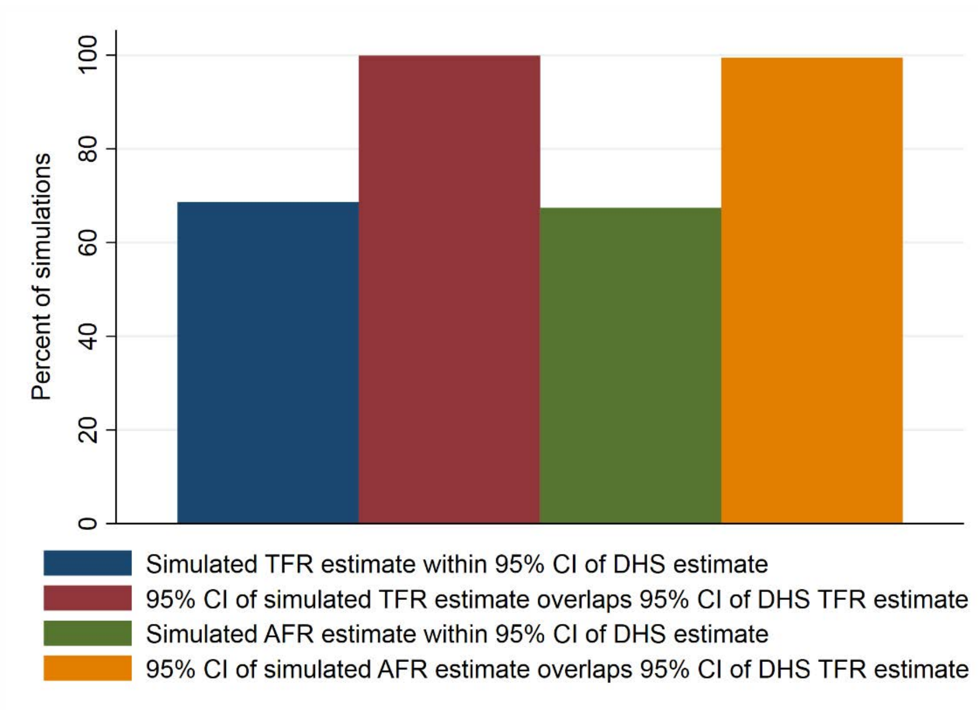



**Figure 4.9. Percentage of 2-year 2-birth fertility estimates from simulated PMA2020 Ethiopia samples that are within confidence bounds of 3-year n-birth estimates, Ethiopia DHS 2016**




As with the prior set of DHS simulations, the percentage of estimates that overlapped actual DHS values was computed. The results are shown in Figure 4.9. Only 69% of TFR point estimates from the simulated subsample fell within the confidence interval of the DHS estimate, although 99.9% of the confidence intervals between the two estimates overlapped. In addition, while only 67% of adolescent fertility rate estimates were within confidence estimates of the DHS estimates, 99.5% produced confidence intervals that overlapped with DHS.

One limitation of this simulation was the sample size of the clusters, which tends to be higher in PMA2020 than in DHS. The simulated subsamples of DHS data produced an unweighted sample size that averaged around 5,800 versus over 7,400 in the PMA2020 survey. A larger sample within each cluster would reduce the standard errors of the estimates and may explain the smaller difference between PMA2020 and DHS standard errors in actuality and those predicted by the simulation. An additional limitation of the sampling technique was that we had to choose from among DHS clusters in a stratum without regard to their 'true' relative size, which may be considered at the survey design stage but is not data available to the end user. While most countries' statistical offices endeavor to classify clusters based on similar size, cluster population may vary slightly. Knowledge of relative cluster size would have produced slightly different samples that may have been closer to DHS values.

# 5. Discussion and Conclusion

DHS Surveys provide many indicators that are used for program planning and monitoring. Some of these are particularly sensitive and noteworthy when the results from a new survey are released. This methodological report focused on six demographic indicators of widespread interest: the Total Fertility Rate (TFR), Infant Mortality Rate (IMR), Under-5 Mortality Rate (U5MR), the Adult Female Mortality Probability (AFMP), the Adult Male Mortality Probability (AMMP) and the Maternal Mortality Ratio (MMR). The IMR, U5MR, AFMP, and AMMP are all probabilities (1q0, 5q0, and 35q15 for women and men, respectively), although the first two are commonly described as rates. Demographers also calculate mortality rates for adult women and men, but here we use the probability of dying between age 15 and 50, partly to be consistent between children and adults and partly because the adult probabilities are more readily available from multiple sources. The MMR is a much more complex indicator that is interpretable as the number of maternal deaths per 100,000 births. DHS calculates the TFR, IMR, and U5MR from the birth histories and the AFMP, AMMP, and MMR from the sibling histories.

It is not unusual for the results of a new survey to be questioned when they differ from some other source. The goal of this report has been to provide guidance on determining whether the estimates of these important indicators are plausible, in the sense of being consistent with other sources, or are not plausible. If the latter conclusion is reached, then further in-depth investigation is required. We do not discuss strategies for such investigations, because they are very specific to the indicator and the survey.

Determining whether a DHS estimate is consistent with other sources is usually a matter of degree. Some differences are expected for a variety of reasons. This report includes a discussion of potential reasons for discrepancies. An important source of variation, even under the best of circumstances, is sampling error. The ramifications of the sample design and sampling error are discussed. Other possible reasons for discrepancies include different reference periods of time. For example, the usual TFR refers to the past 3 years, and is centered 1.5 years before the mean date of interview. The IMR and U5MR refer to the past 5 years, and are centered 2.5 years before the mean date of interview. The AFMP, AMMP, and MMR refer to the past 7 years, and are centered 3.5 years before the mean date of interview. When comparing sources, it is important to consider the reference period. Differences can also arise from the use of indirect versus direct estimates (DHS estimates are always direct), the adjustment of data (DHS does not make adjustments), and the use of complete versus truncated birth histories (DHS birth histories are always complete). All comparisons must consider these differences.

Chapter 3 of the report analyzed 51 surveys conducted since 2010 in terms of the estimates of the TFR, IMR, and U5MR; 31 of those surveys included sibling histories, in terms of the estimates of the AFMP, AMMP, and MMR. These estimates were systematically compared with estimates of the TFR, AFMP, and AMMP from the UN Population Division, estimates of the IMR and U5MR from the UN's Inter Agency Group for Child Mortality Estimation (IGME), and MMR estimates from the World Health Organization (WHO). These official, adjusted estimates were selected for the dates or time intervals closest to the reference dates for the corresponding DHS estimates. All estimates, from both DHS and the UN agencies, were downloaded from publicly accessible websites.

To provide more focus, the analysis centered on the Afghanistan 2015 and Angola 2015-16 surveys. A set of 21 figures guided the comparison of the DHS estimates and the UN estimates. The structure of the comparisons could be applied to other sources, although it is always necessary to consider the potential kinds of differences discussed earlier. The UN estimates are not independent of the DHS estimates, because DHS data are among the many ingredients of the UN estimates, but they have been carefully adjusted with statistical and demographic models. The figures used for the comparisons are accompanied by tables that provide specific numerical values, and in many cases, confidence intervals or uncertainty intervals. The findings from these comparisons will not be repeated here, because they were solely intended to be illustrative of the strategy used for making the comparisons.

In countries that lack comprehensive and accurate vital registration, surveys estimate vital rates through retrospective data. Differing fertility estimates produced by DHS and non-DHS surveys such as MICS and PMA2020 that conduct fieldwork within the same timeframe may cause confusion in interpretation and trend analysis. Chapter 4 compared DHS fertility rate estimates with those from an alternative source, PMA2020, explored the effect of a slightly different methodology for computing rates on DHS estimates, and simulated subsamples of Ethiopia 2016 DHS data that are similar to PMA2020's Round 4 Ethiopia survey to assess the effect of a different sampling strategy and rate estimation method on fertility rate results.

Chapter 4 shows that PMA2020 fertility rates estimated with a standard methodology have a range of 5% to 22% difference in TFRs and a 4% to 17% difference in adolescent fertility rates compared to those published by The DHS Program. The standard errors in PMA2020 fertility estimates also range from 65% to 222% higher than corresponding DHS estimates. Beyond survey factors such as questionnaire design and date imputation, these estimates are produced by techniques that slightly differ in methodology. DHS publishes 3-year n-birth fertility rates while published PMA2020 fertility estimates tend to be 2-year 2-birth adjusted rates.[6] In addition, survey design effects (sample frames and sizes) may play a role in differences. In Ethiopia, for example, both surveys used a two-stage stratified random sample, although PMA2020 grouped several regions together as one, sampled a smaller number of clusters in each region, and sampled a larger number of women per cluster.

To assess the effect of alternate measurement on fertility rates from the exact same survey, Chapter 4 used all 256 available standard, continuous, and interim DHS datasets to compare the results of a 2-year 2-birth adjusted estimation technique versus a 3-year n-birth technique with the same data. Our results show that fertility rates calculated with a 2-year 2-birth adjusted fertility rate method should typically average only a one percentage point difference in total fertility and adolescent fertility rates. However, there may be as much as a 7 percentage point difference in total fertility rates and a 29 percentage point in adolescent fertility rates produced by these two different methods when using the same data.

In addition to the differences produced by alternate techniques of fertility measurement, it is important to consider the overall design effect of a different sample. In Ethiopia, we simulated 1,000 subsamples of 2016 DHS data with a cluster distribution similar to PMA2020 and computed rates with a 2-year 2-birth adjusted estimation technique. We found that these combined differences in sampling and methodology typically produce an average 3% to 4% difference but may plausibly produce as much as a 10% difference in TFR, a 23% difference in adolescent fertility rate in Ethiopia, and typically around double the standard errors of DHS estimates. Notably, since the PMA2020 cluster sample sizes are larger than those used by DHS, this simulation is likely to overestimate the effect of sampling difference on results from Ethiopia.

When comparing fertility estimates from different surveys, it is important to recognize that differences in sampling, sample size, and overall fertility rate computation methods may—even with the same data—drive apparent differences in fertility rates. However, we find that measurement and sampling factors on the same survey data typically produce results that have overlapping confidence intervals. When results from alternate surveys are unusually different from DHS, it may also be worth considering differences in the survey timeframe, wording of questionnaires, and date imputation that could help account for the observed difference. Such factors are beyond what we can assess from DHS data alone. A recent report by Choi et al. finds that random redistribution of excess January births in PMA2020 data would increase TFR estimates by an average of 3%. Surveys such as MICS, which may employ a truncated birth history, base

[6] Authors' correspondence with PMA2020 staff, August 2017.

fertility on a different reference period, or do not employ a PMA2020-type adjustment for multiple births should also be expected to produce variable, and lower fertility rates than the published DHS estimates.

# References

Ahmed, S., Q. Li, C. Scrafford, and T.W. Pullum. 2014. *An Assessment of DHS Maternal Mortality Data and Estimates*. DHS Methodological Reports No. 13. Rockville, Maryland, USA: ICF International. Available at http://dhsprogram.com/pubs/pdf/MR13/MR13.pdf.

Alkema, L., and D. You. 2012. "Child Mortality Estimation: A Comparison of UN IGME and IHME Estimates of Levels and Trends in Under-Five Mortality Rates and Deaths." *PLoS Medicine* 9(8): e1001288.

Brady, E., and K. Hill. 2017. "Testing Survey-Based Methods for Rapid Monitoring of Child Mortality, with Implications for Summary Birth History Data." *PloS One* 12(4): e0176366.

Choi, Y., Q. Li, and B. Zachary. 2017. Fertility Data in Performance Monitoring and Accountability 2020 Surveys. *Performance Monitoring and Accountability 2020 Methodological Reports No. 2.* Baltimore, Maryland, USA: Bill & Melinda Gates Institute for Population and Reproductive Health, Johns Hopkins University Bloomberg School of Public Health.

The DHS Program. 2017. *STATcompiler*. The DHS Program. Available at http://statcompiler.com/en/.

Hancioglu, A., and F. Arnold. 2013. "Measuring Coverage in MNCH: Tracking Progress in Health for Women and Children Using DHS and MICS Household Surveys." *PLoS Medicine* 10(5): e1001391.

Masquelier, B., G. Reniers, and G. Pison. 2014. "Divergences in Trends in Child and Adult Mortality in Sub-Saharan Africa: Survey Evidence on the Survival of Children and Siblings." *Population Studies* 68(2): 161-177.

Pullum, T.W., and S. Becker. 2014. *Evidence of Omission and Displacement in DHS Birth Histories*. DHS Methodological Reports No. 11. Rockville, Maryland, USA: ICF International. Available at http://dhsprogram.com/pubs/pdf/MR11/MR11.pdf.

Pullum, T.W., and S. 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.

Schoumaker, B. 2013. "A Stata Module for Computing Fertility Rates and TFRs from Birth Histories: TFR2." *Demographic Research* 28:1093-1144.

Schoumaker, B. 2014. *Quality and Consistency of DHS Fertility Estimates, 1990 to 2012*. DHS Methodological Reports No. 12. Rockville, Maryland, USA: ICF International. Available at http://dhsprogram.com/pubs/pdf/MR12/MR12.pdf.

Silva, R. 2012. "Child Mortality Estimation: Consistency of Under-Five Mortality Rate Estimates Using Full Birth Histories and Summary Birth Histories." *PLoS Medicine* 9(8): e1001296.

UNICEF. 2017. *Child Survival*. Available at https://data.unicef.org/topic/child-survival/.

United Nations. 2017. *World Population Prospects 2017*. United Nations DESA Population Division. Available at https://esa.un.org/unpd/wpp/.

United Nations. 1983. *Manual 10-Indirect Techniques for Demographic Estimation*. New York, USA: UN Department of International Economic and Social Affairs.

WHO. 2015. *Trends in Maternal Mortality: 1990 to 2015*. Geneva, Switzerland: WHO. Available at http://www.who.int/reproductivehealth/publications/monitoring/maternal-mortality-2015/en/.

Zimmerman, L., H. Olson, A. Tsui, and S. Radloff. 2017. "PMA2020: Rapid Turn-around Survey Data to Monitor Family Planning Service and Practice in Ten Countries." *Studies in Family Planning*.

# Appendices

**Table A.1. List of surveys included in this report and reference dates for the rates**

| Survey | Mean Date of Interview | Reference date: TFR | Reference date: Child mortality | Reference date: Adult Mortality |
|---|---|---|---|---|
| Afghanistan 2015 | 2015.8 | 2014.3 | 2013.3 | 2012.3 |
| Angola 2015-16 | 2016.0 | 2014.5 | 2013.5 | 2012.5 |
| Armenia 2010 | 2010.9 | 2009.4 | 2008.4 | 2007.4 |
| Bangladesh 2014 | 2014.7 | 2013.2 | 2012.2 | 2011.2 |
| Benin 2011-12 | 2012.1 | 2010.6 | 2009.6 | 2008.6 |
| Burkina Faso 2010 | 2010.7 | 2009.2 | 2008.2 | 2007.2 |
| Burundi 2010 | 2010.8 | 2009.3 | 2008.3 | 2007.3 |
| Cambodia 2014 | 2014.6 | 2013.1 | 2012.1 | 2011.1 |
| Cameroon 2011 | 2011.3 | 2009.8 | 2008.8 | 2007.8 |
| Chad 2014-15 | 2015.1 | 2013.6 | 2012.6 | 2011.6 |
| Colombia 2015 | 2015.6 | 2014.1 | 2013.1 | 2012.1 |
| Comoros 2012 | 2012.8 | 2011.3 | 2010.3 | 2009.3 |
| Congo 2011-12 | 2011.9 | 2010.4 | 2009.4 | 2008.4 |
| Cote d'Ivoire 2011-12 | 2012.1 | 2010.6 | 2009.6 | 2008.6 |
| Dominican Republic 2013 | 2013.6 | 2012.1 | 2011.1 | 2010.1 |
| DRC 2013-14 | 2013.9 | 2012.4 | 2011.4 | 2010.4 |
| Egypt 2014 | 2014.4 | 2012.9 | 2011.9 | 2010.9 |
| Ethiopia 2011 | 2011.2 | 2009.7 | 2008.7 | 2007.7 |
| Gabon 2012 | 2012.2 | 2010.7 | 2009.7 | 2008.7 |
| Gambia 2013 | 2013.2 | 2011.7 | 2010.7 | 2009.7 |
| Ghana 2014 | 2014.8 | 2013.3 | 2012.3 | 2011.3 |
| Guatemala 2014-15 | 2015.2 | 2013.7 | 2012.7 | 2011.7 |
| Guinea 2012 | 2012.6 | 2011.1 | 2010.1 | 2009.1 |
| Haiti 2012 | 2012.3 | 2010.8 | 2009.8 | 2008.8 |
| Honduras 2011-12 | 2012.1 | 2010.6 | 2009.6 | 2008.6 |
| Indonesia 2012 | 2012.4 | 2010.9 | 2009.9 | 2008.9 |
| Jordan 2012 | 2012.8 | 2011.3 | 2010.3 | 2009.3 |
| Kenya 2014 | 2014.6 | 2013.1 | 2012.1 | 2011.1 |
| Kyrgyz Republic 2012 | 2012.8 | 2011.3 | 2010.3 | 2009.3 |
| Lesotho 2014 | 2014.8 | 2013.3 | 2012.3 | 2011.3 |
| Liberia 2013 | 2013.4 | 2011.9 | 2010.9 | 2009.9 |
| Malawi 2015-16 | 2015.9 | 2014.4 | 2013.4 | 2012.4 |
| Mali 2012-13 | 2013.0 | 2011.5 | 2010.5 | 2009.5 |
| Mozambique 2011 | 2011.6 | 2010.1 | 2009.1 | 2008.1 |
| Myanmar 2015-16 | 2016.1 | 2014.6 | 2013.6 | 2012.6 |
| Namibia 2013 | 2013.6 | 2012.1 | 2011.1 | 2010.1 |
| Nepal 2011 | 2011.3 | 2009.8 | 2008.8 | 2007.8 |
| Niger 2012 | 2012.3 | 2010.8 | 2009.8 | 2008.8 |
| Nigeria 2013 | 2013.3 | 2011.8 | 2010.8 | 2009.8 |
| Pakistan 2012-13 | 2012.9 | 2011.4 | 2010.4 | 2009.4 |
| Philippines 2013 | 2013.7 | 2012.2 | 2011.2 | 2010.2 |
| Rwanda 2014-15 | 2015.1 | 2013.6 | 2012.6 | 2011.6 |
| Senegal 2010-11 | 2011.0 | 2009.5 | 2008.5 | 2007.5 |
| Sierra Leone 2013 | 2013.6 | 2012.1 | 2011.1 | 2010.1 |
| Tajikistan 2012 | 2012.6 | 2011.1 | 2010.1 | 2009.1 |
| Tanzania 2015-16 | 2015.9 | 2014.4 | 2013.4 | 2012.4 |
| Togo 2013-14 | 2014.1 | 2012.6 | 2011.6 | 2010.6 |
| Uganda 2011 | 2011.7 | 2010.2 | 2009.2 | 2008.2 |
| Yemen 2013 | 2013.8 | 2012.3 | 2011.3 | 2010.3 |
| Zambia 2013-14 | 2013.9 | 2012.4 | 2011.4 | 2010.4 |
| Zimbabwe 2015 | 2015.7 | 2014.2 | 2013.2 | 2012.2 |

**Figure A.1. TFR with 95% confidence intervals for 51 DHS surveys and the UN estimate for each survey**

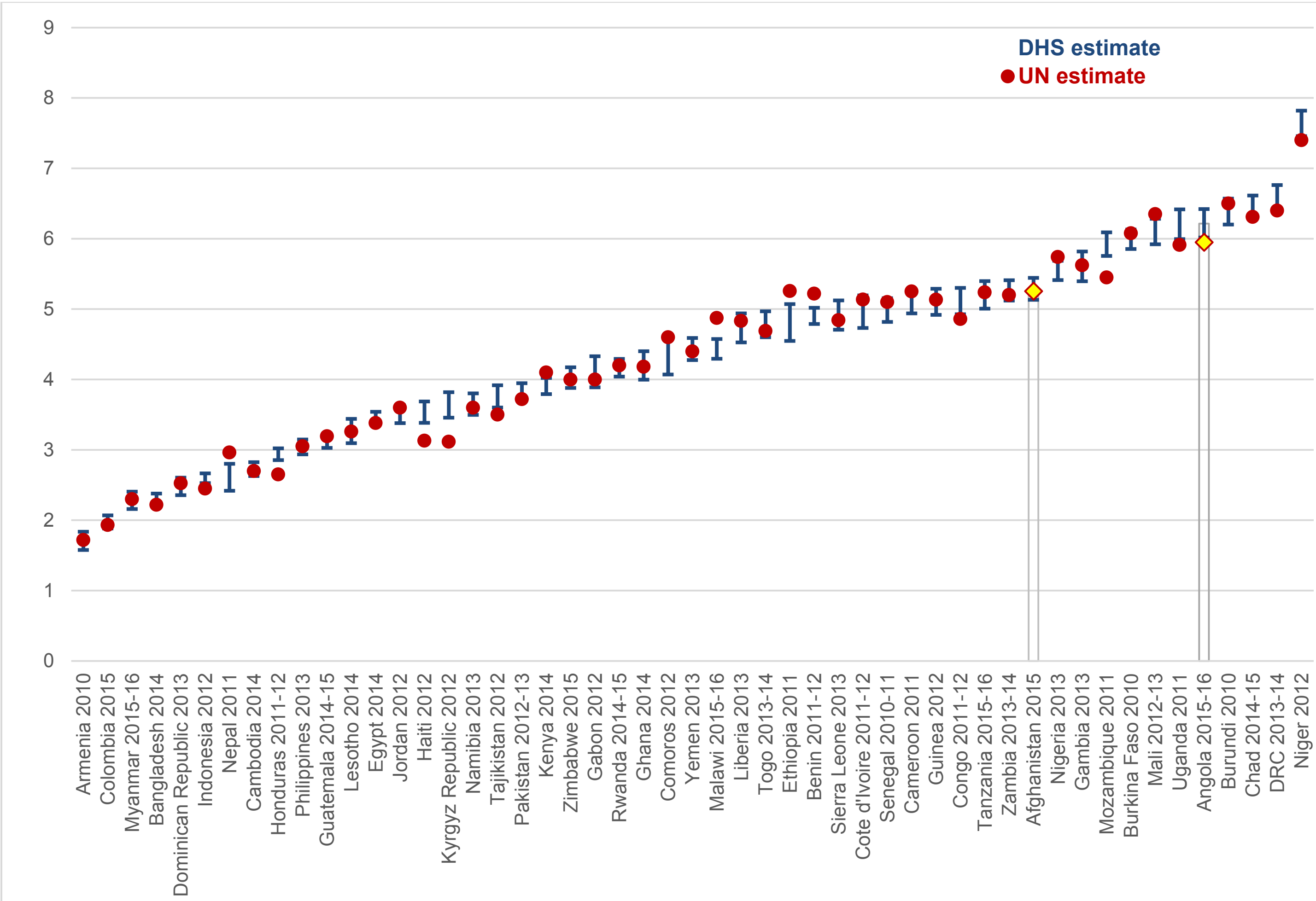


**Figure A.2. TFR estimates with difference between DHS and UN estimate**

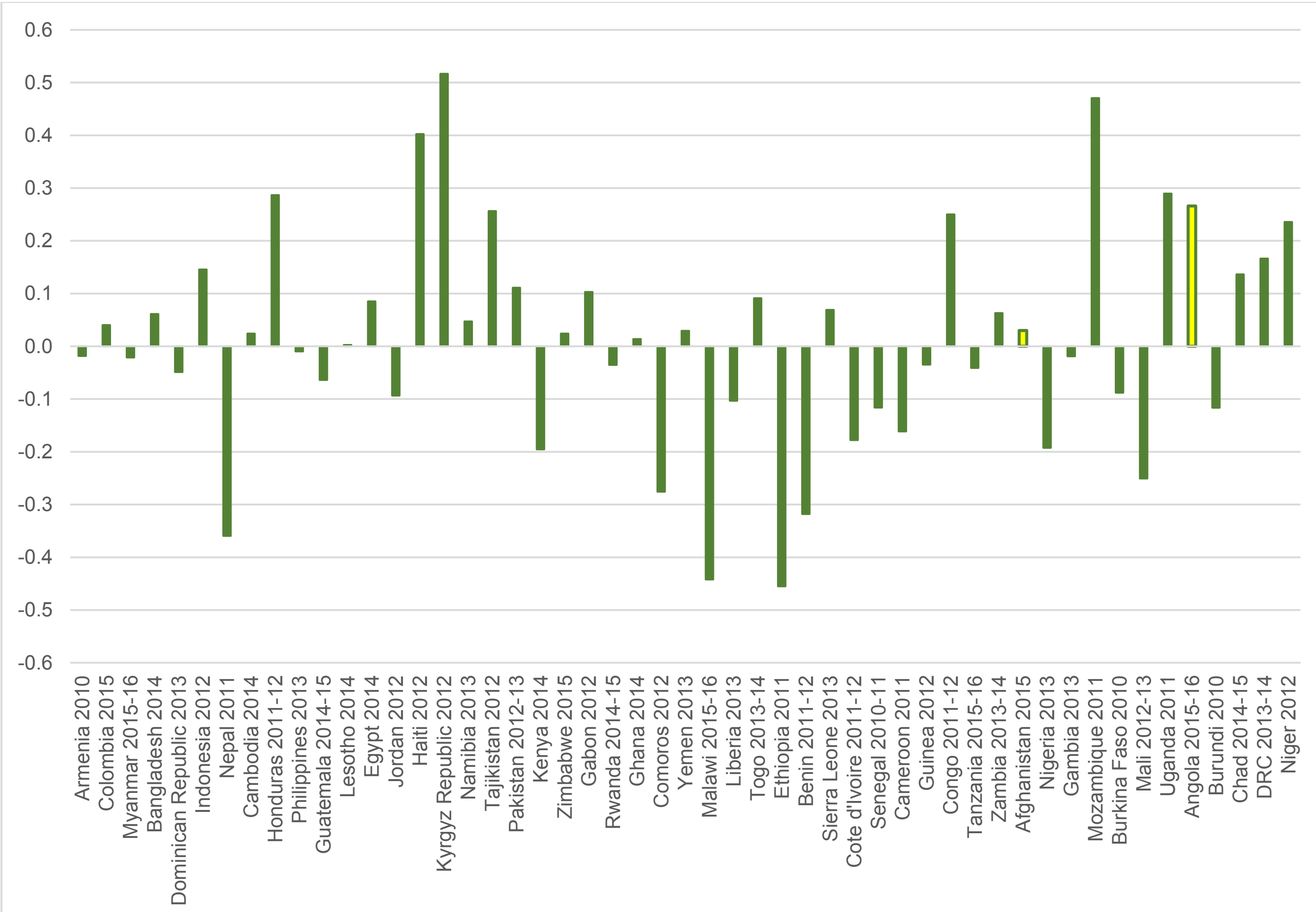


**Figure A.3. IMR DHS compared to IGME**

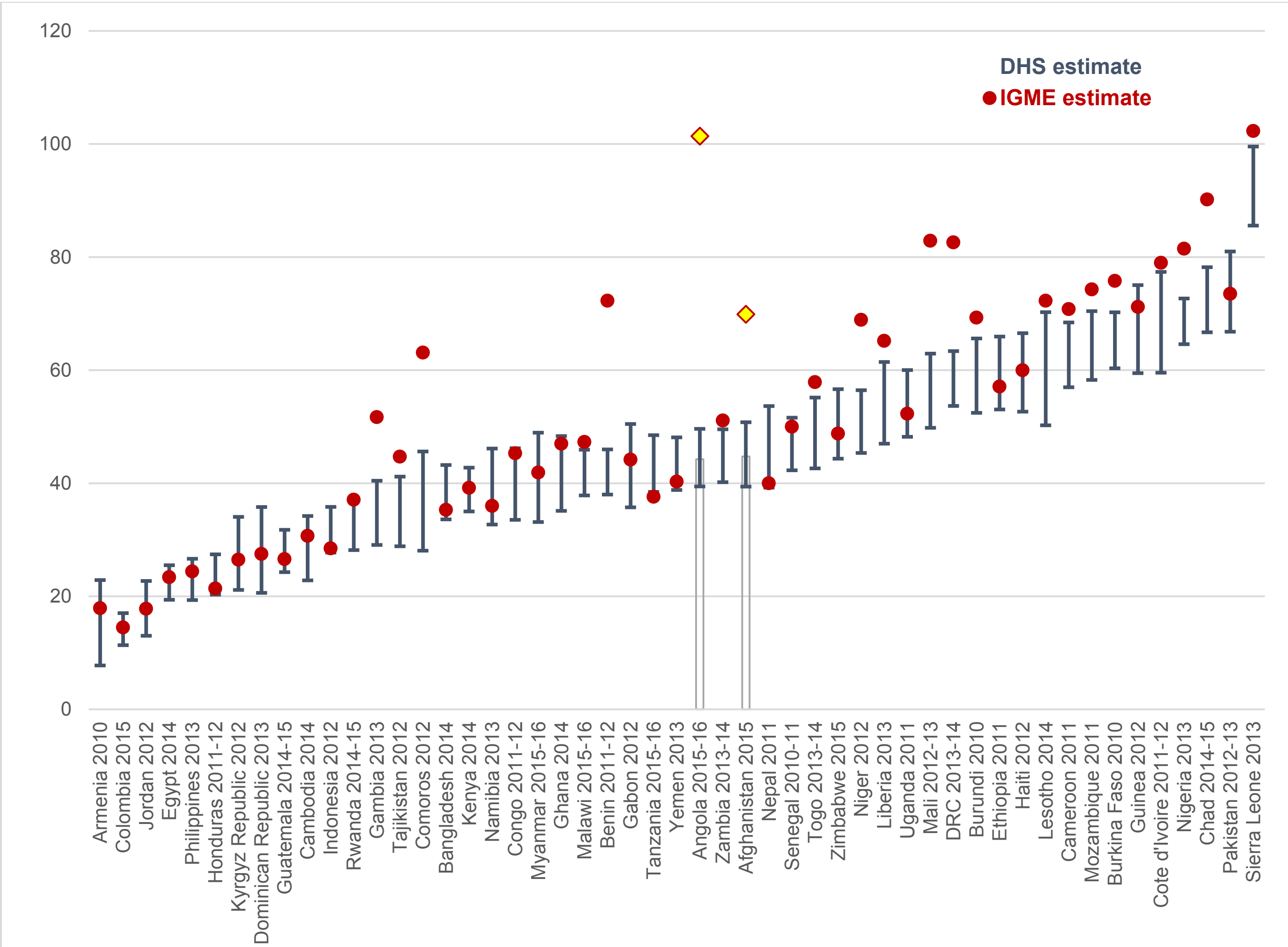


**Figure A.4. IMR for DHS compared to IGME with differences**

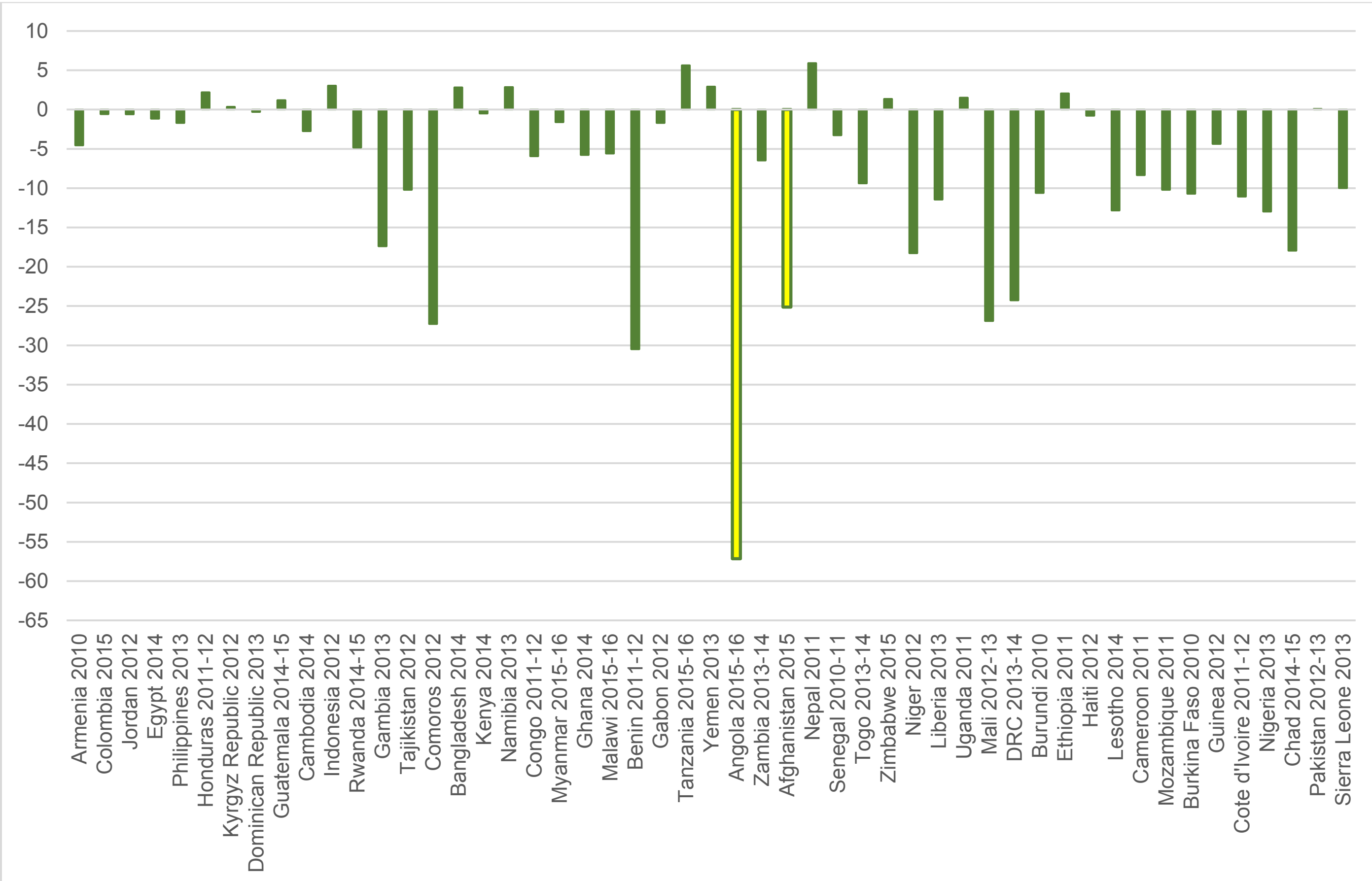


**Figure A.5. U5MR DHS compared to IGME**

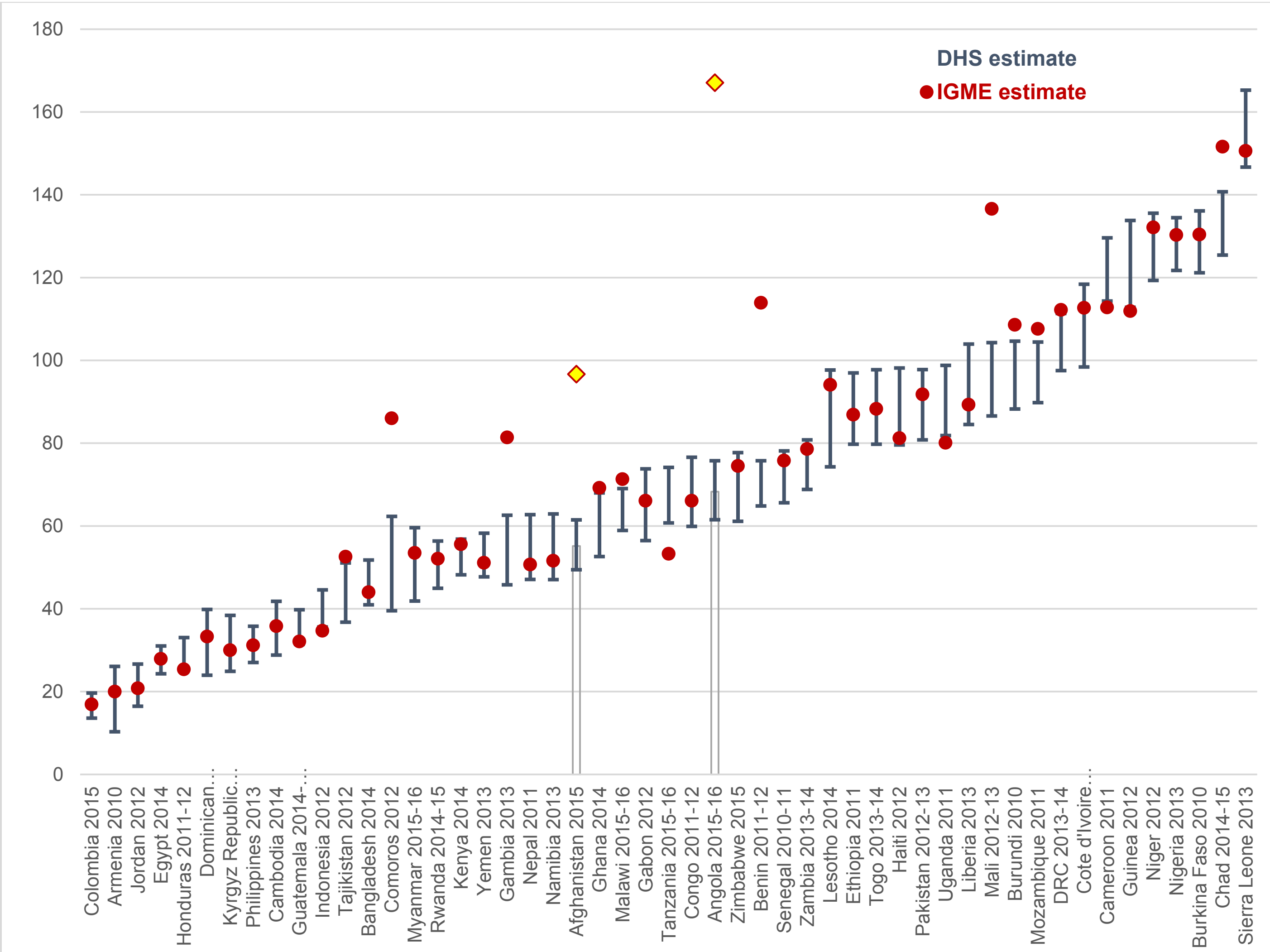


**Figure A.6. U5MR DHS compared to IGME with differences**

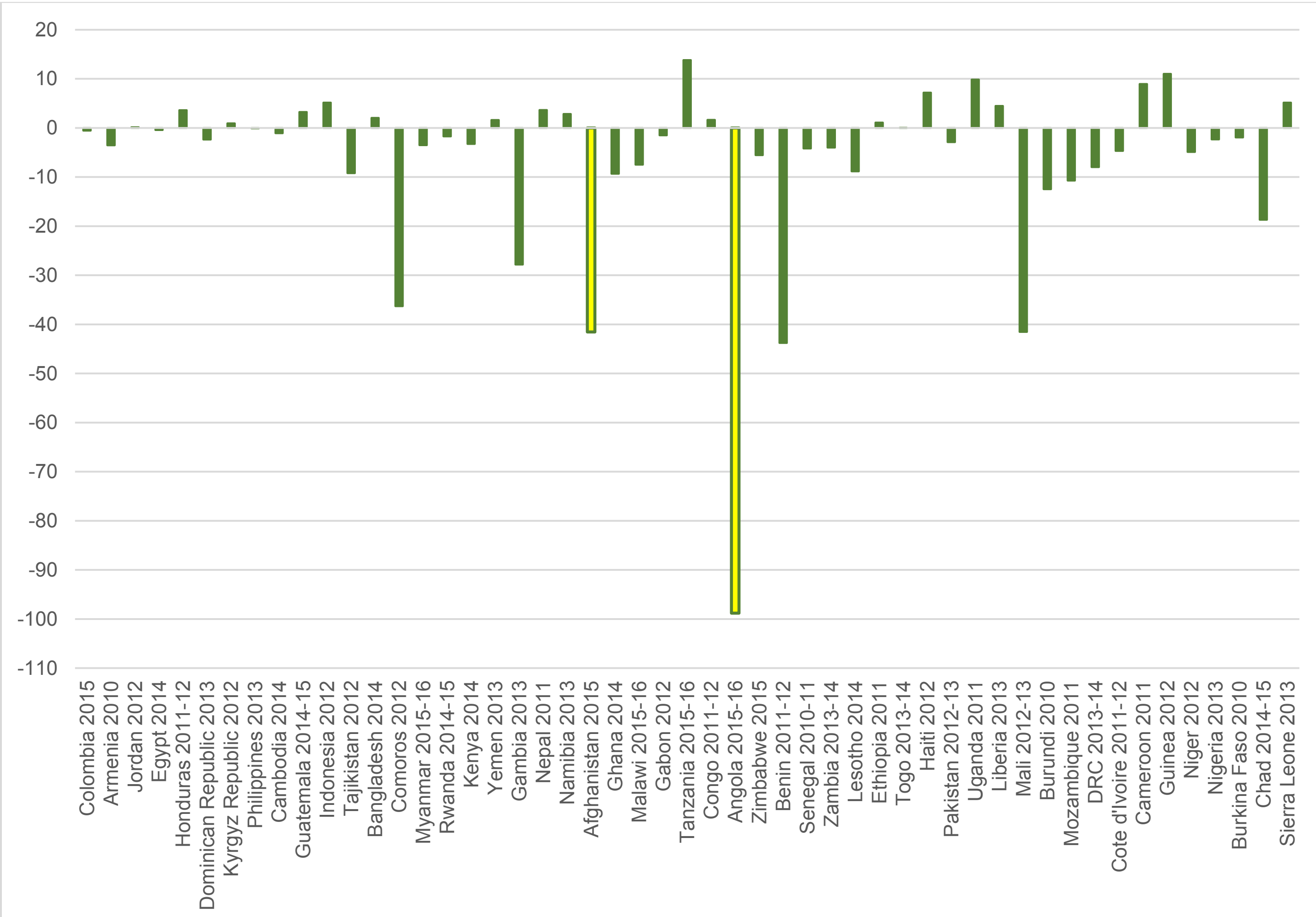


**Figure A.7. IMR versus U5MR using DHS and IGME estimates**

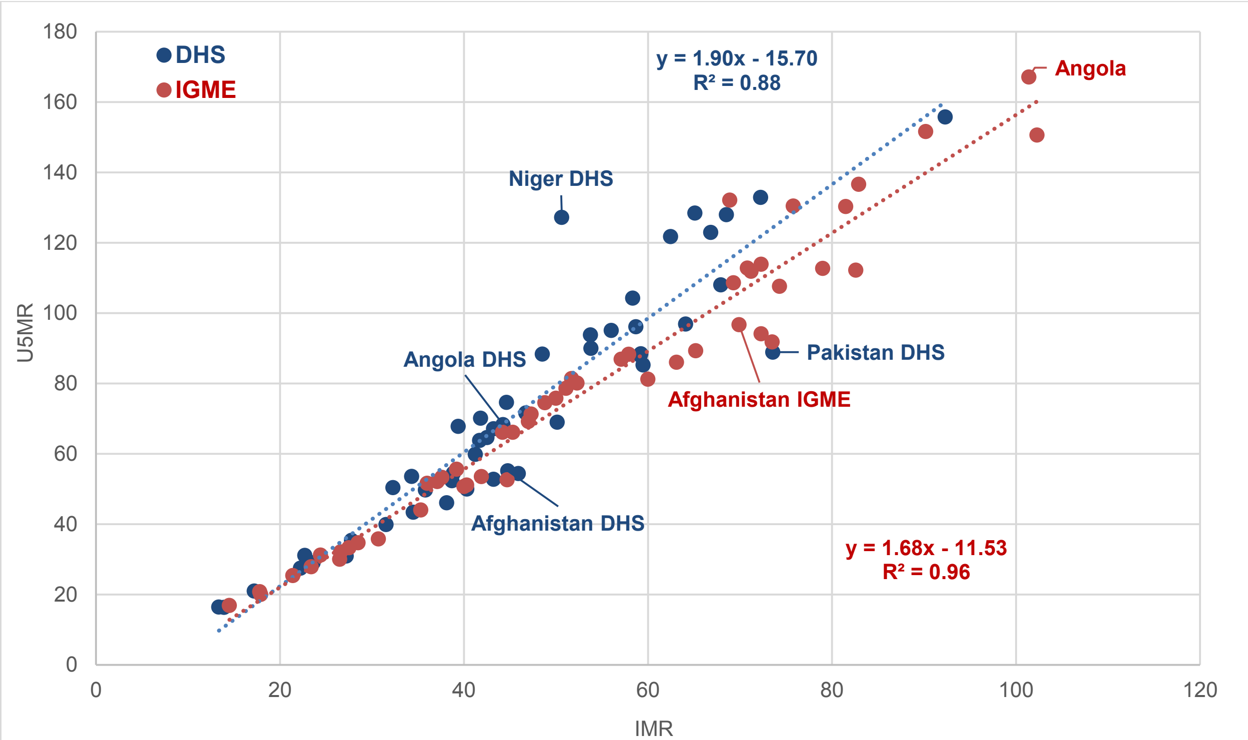



**Figure A.8. Log of IMR versus log of U5MR using DHS and IGME estimates**

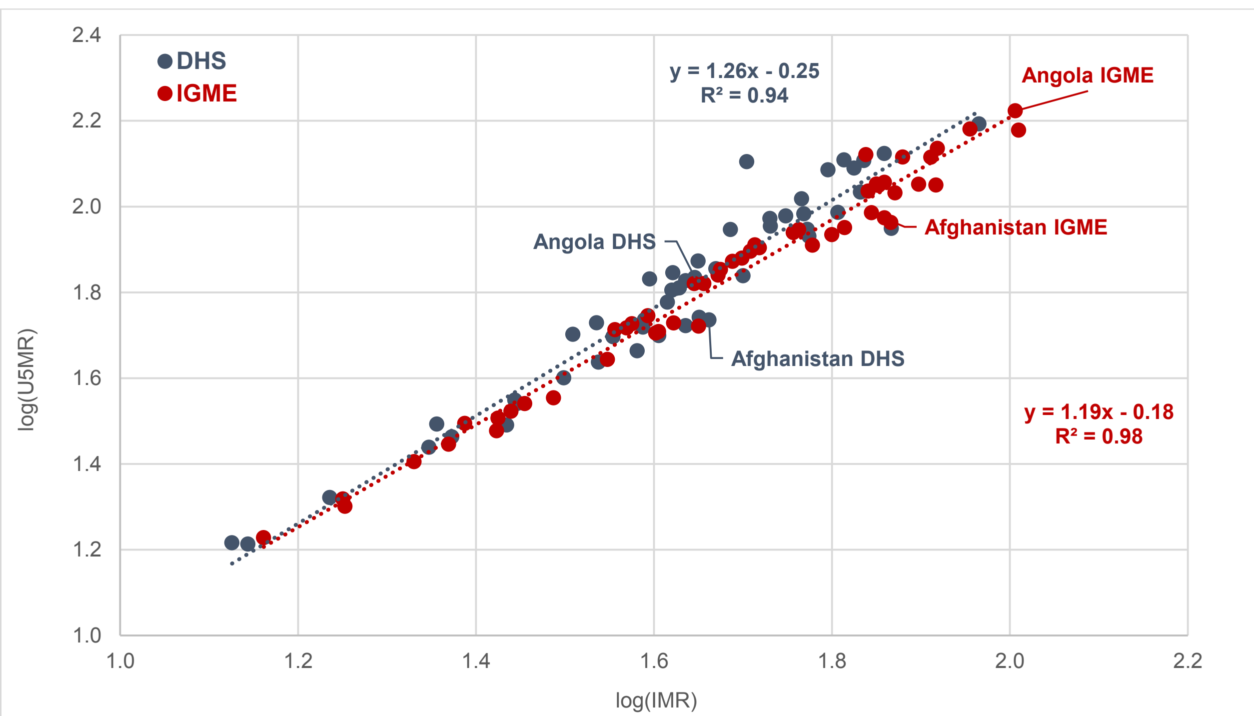


**Figure A.9. Log of the ratio of DHS and IGME estimates for IMR versus U5MR**

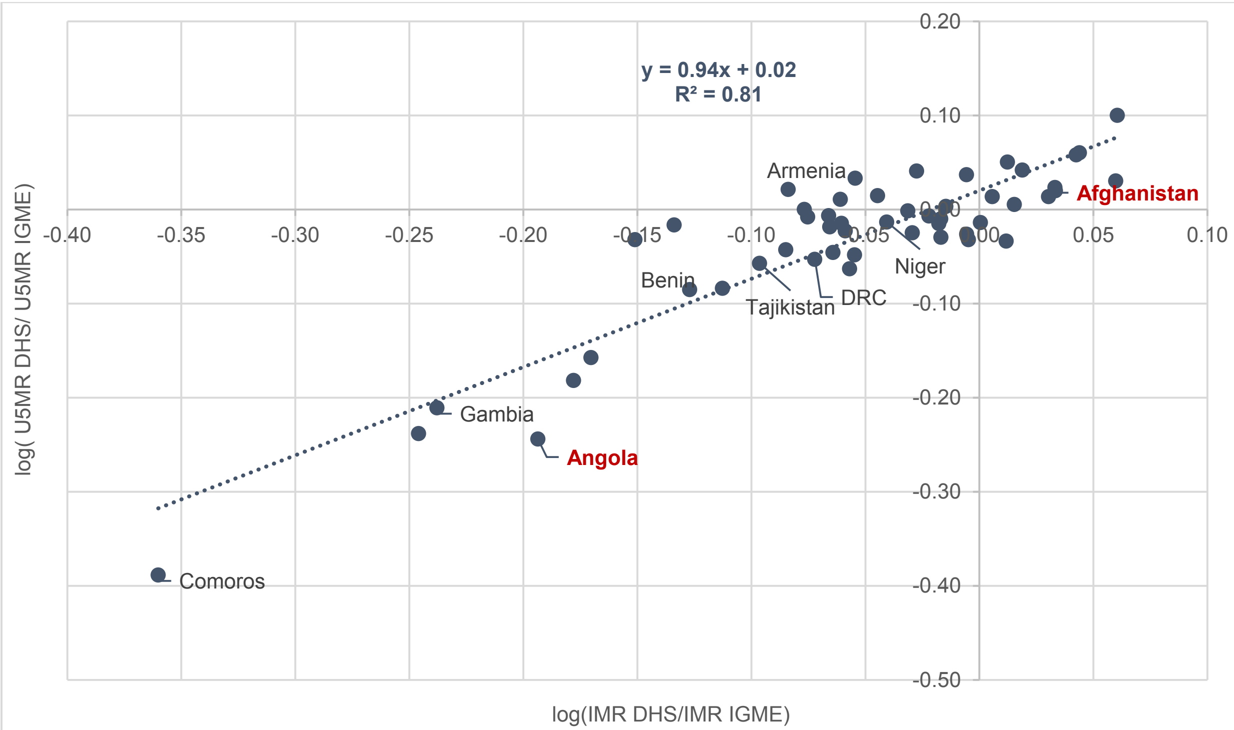



**Figure A.10. Female probability of dying between age 15 and 50**

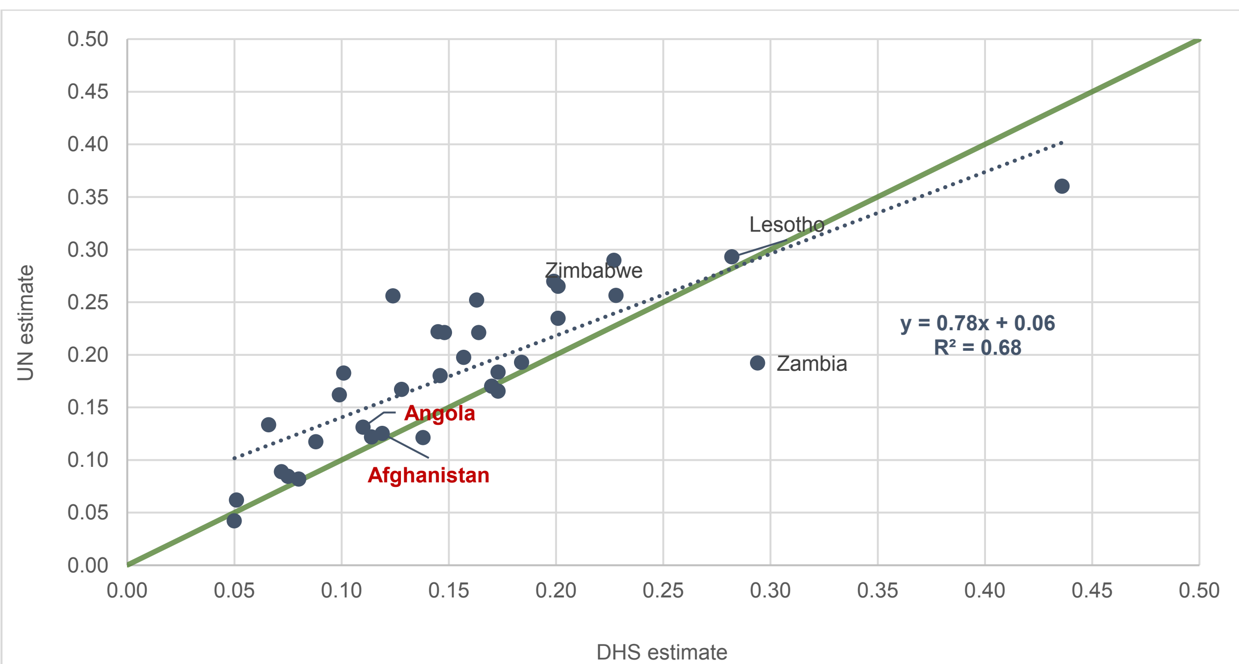


**Figure A.11. Male probability of dying between age 15 and 50**

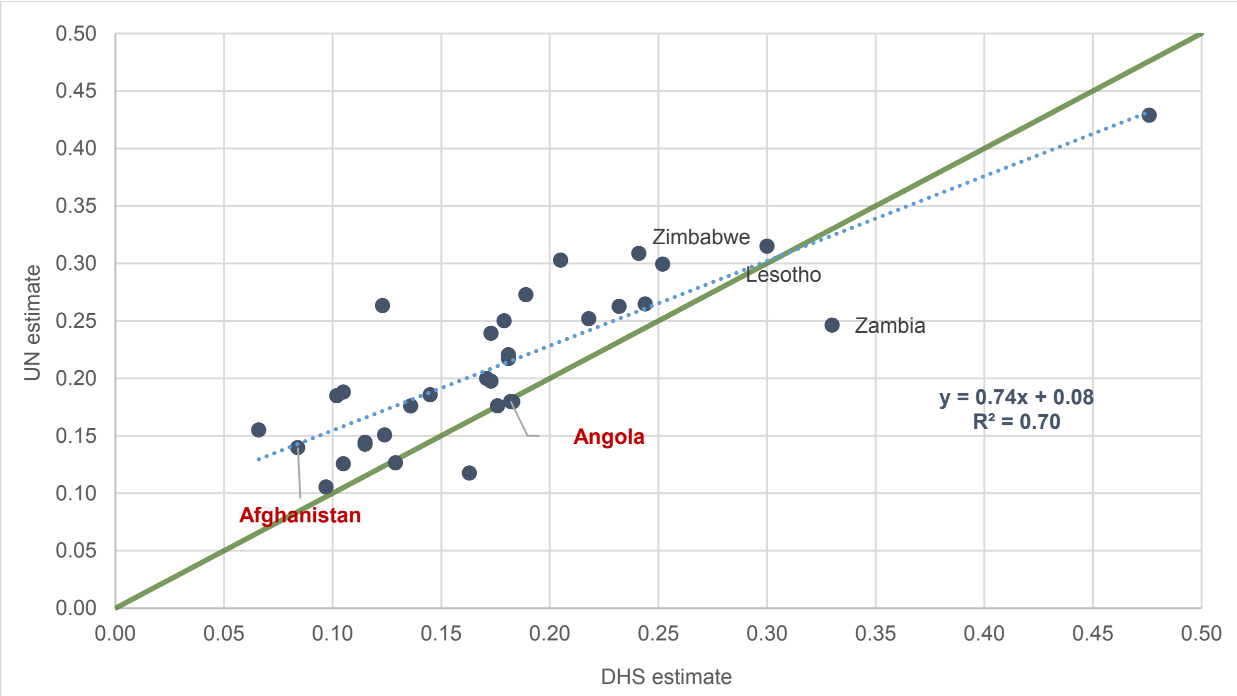



**Figure A.12. Male versus Female Adult Mortality**

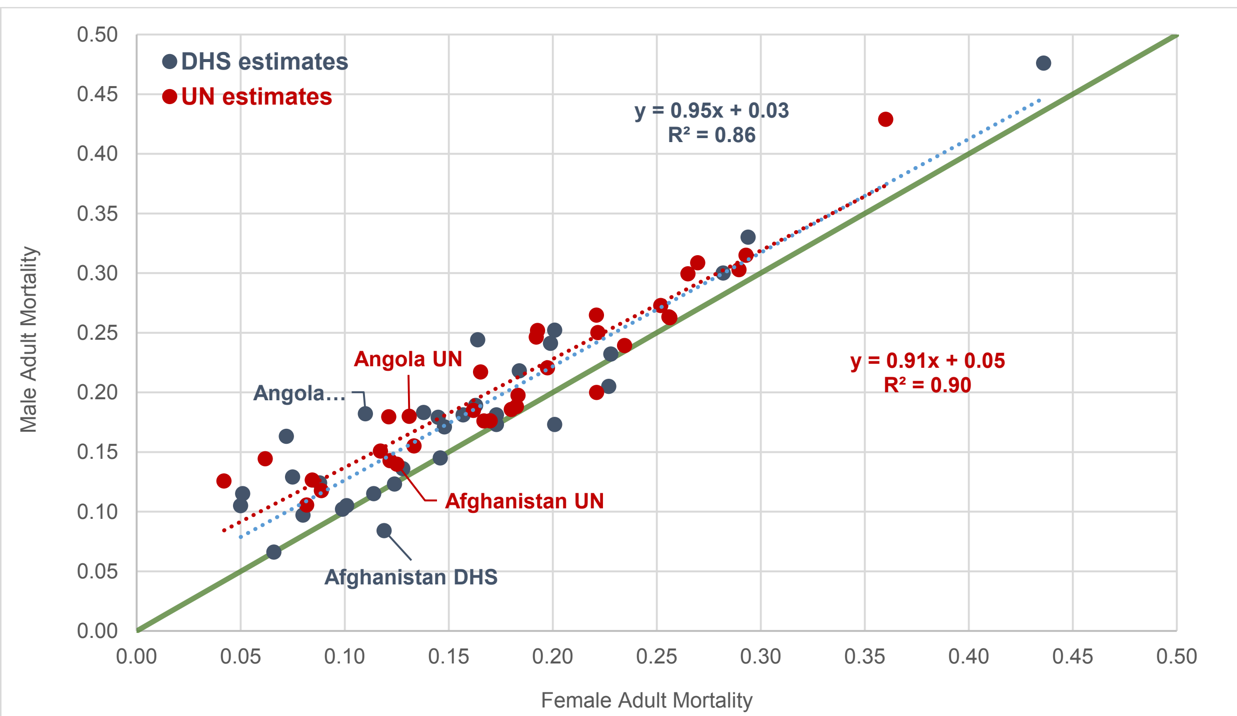


**Figure A.13. Difference in Male Adult Mortality versus Difference in Female Adult Mortality**

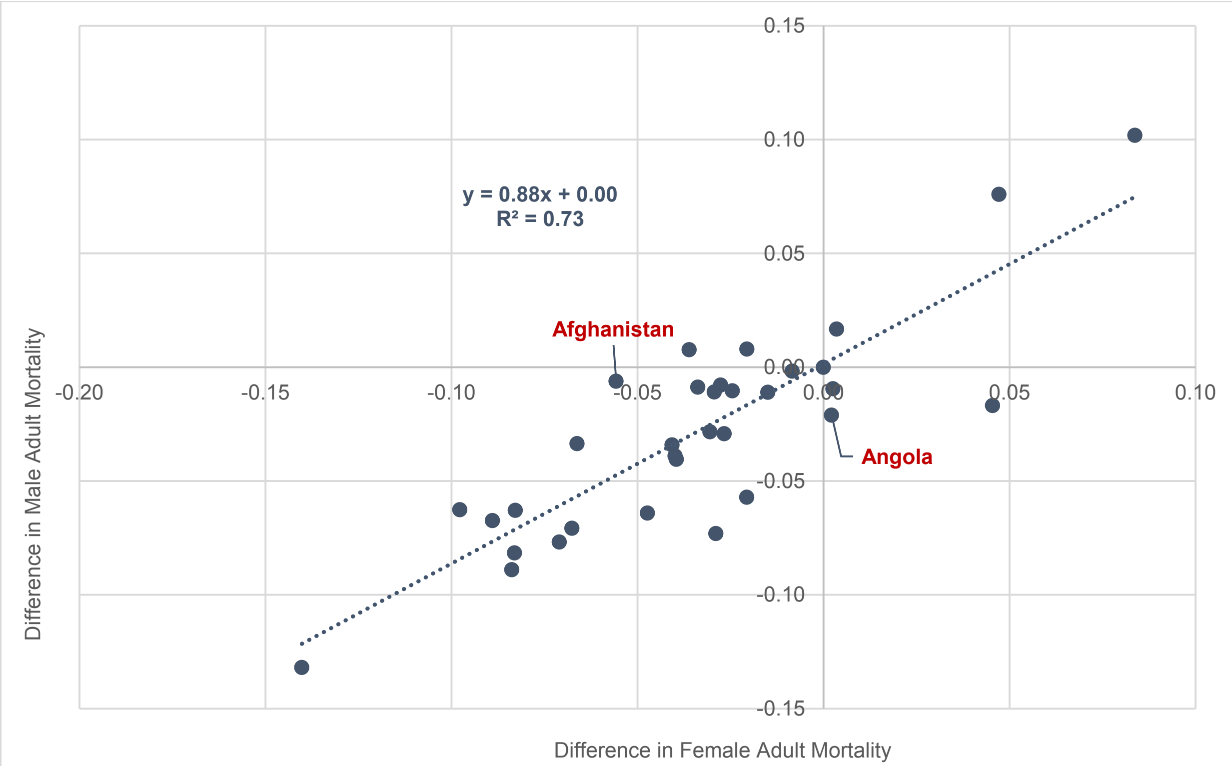



**Figure A.14. Adult Female Mortality versus U5MR**

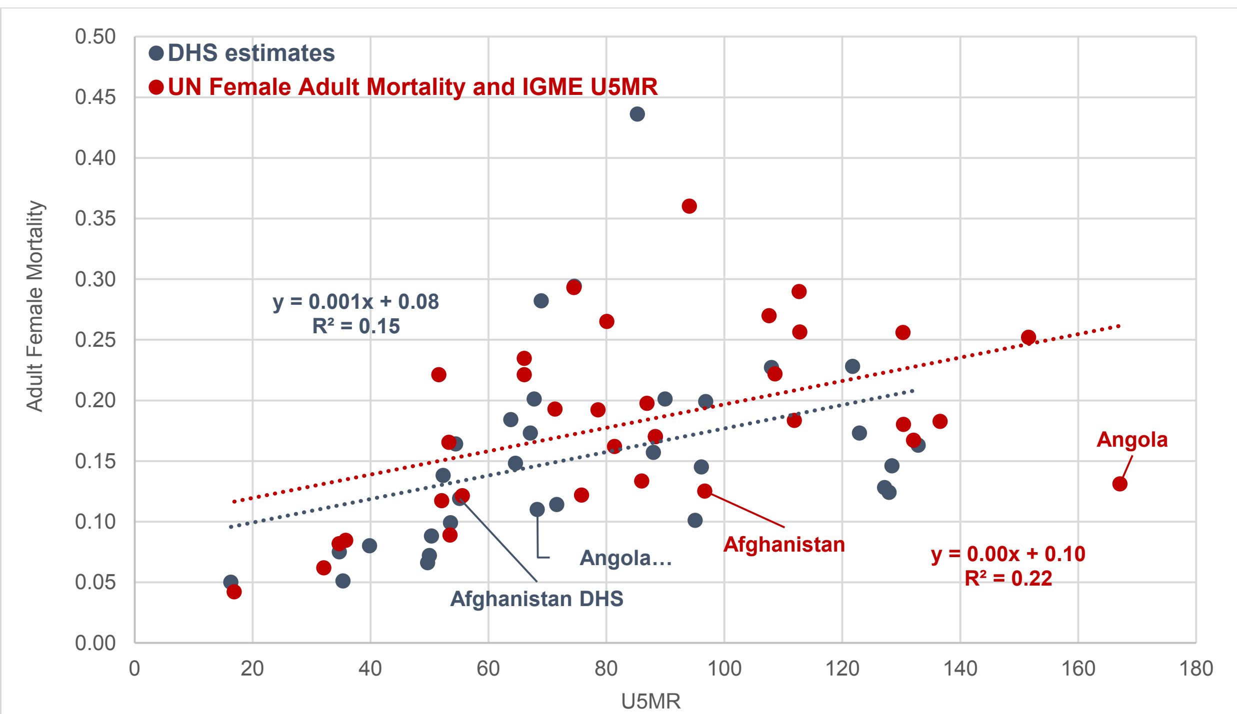


**Figure A.15. Difference in Adult Female Mortality versus Difference in U5MR**

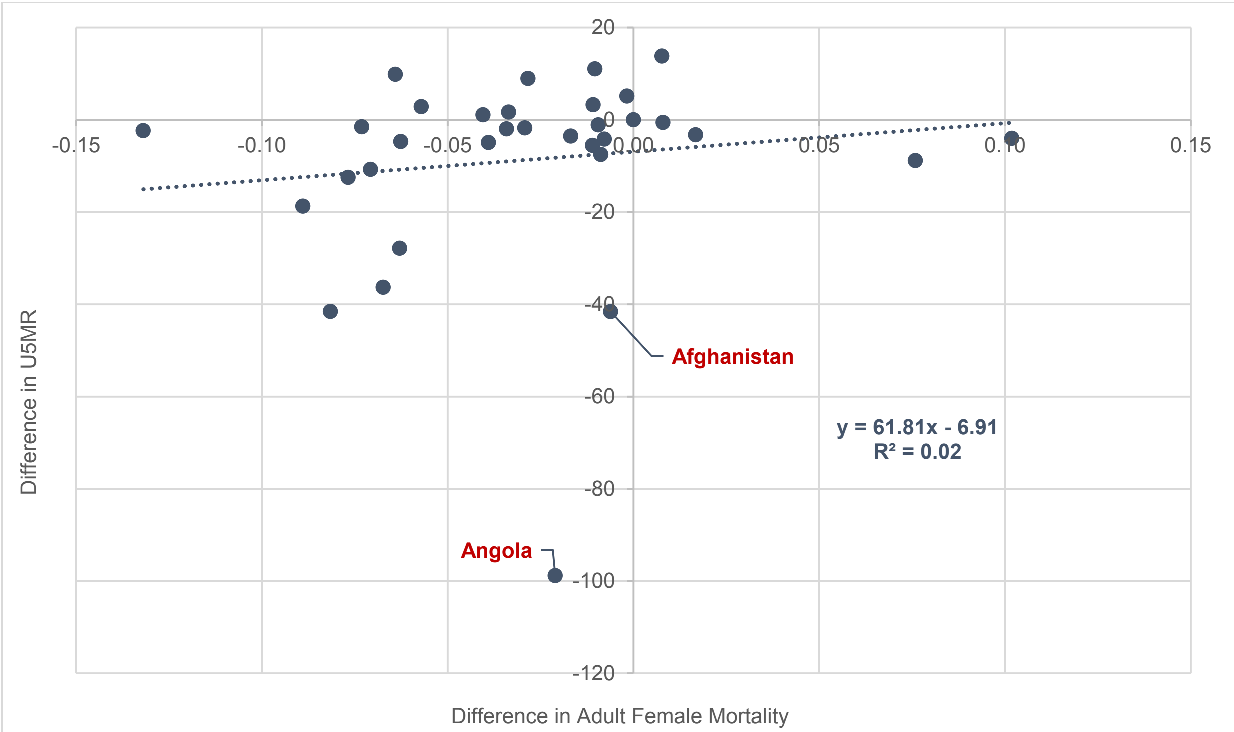


**Figure A.16. MMR using DHS and WHO estimates**

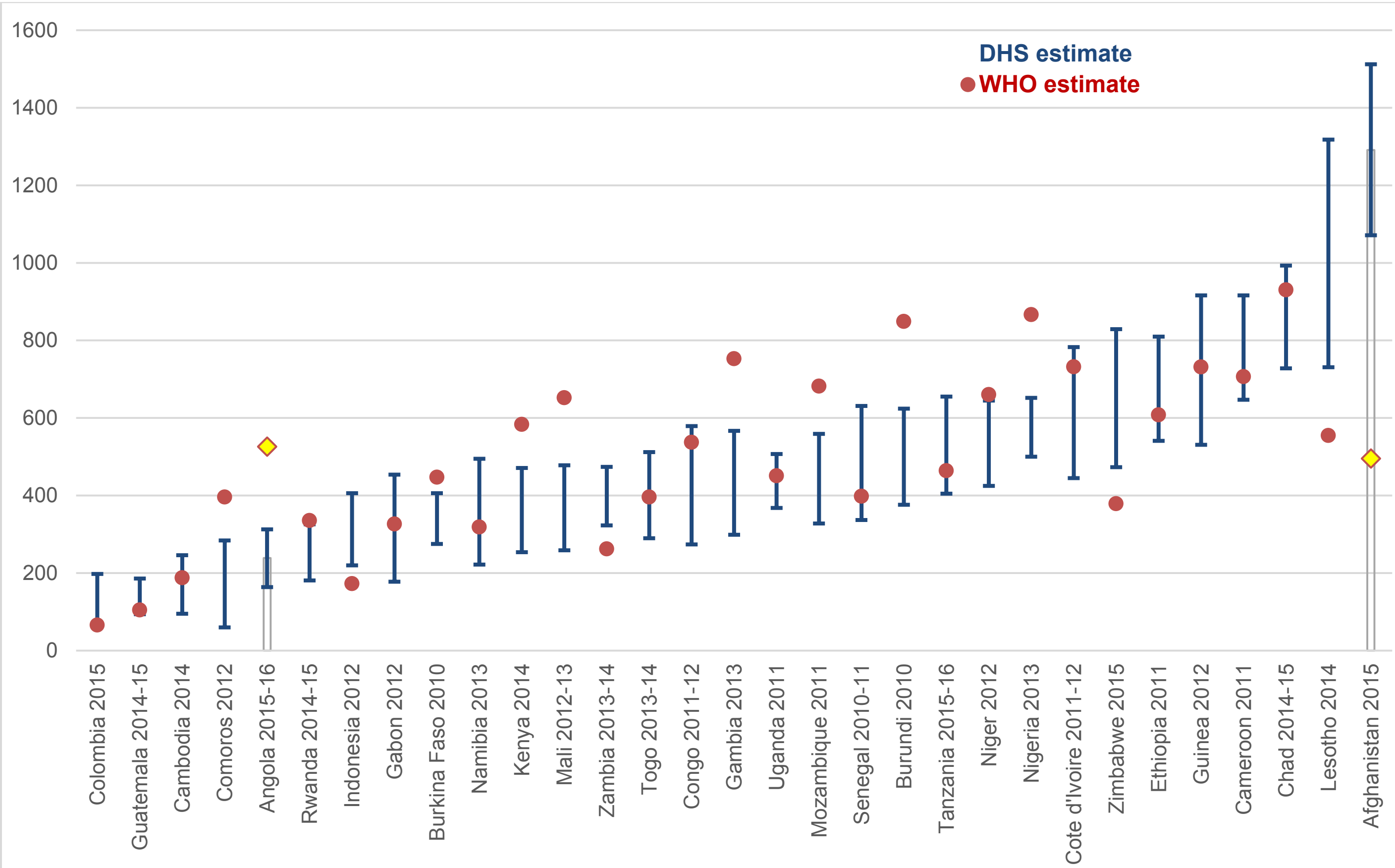


**Figure A.17. MMR using DHS and WHO estimates and the difference between DHS and WHO**

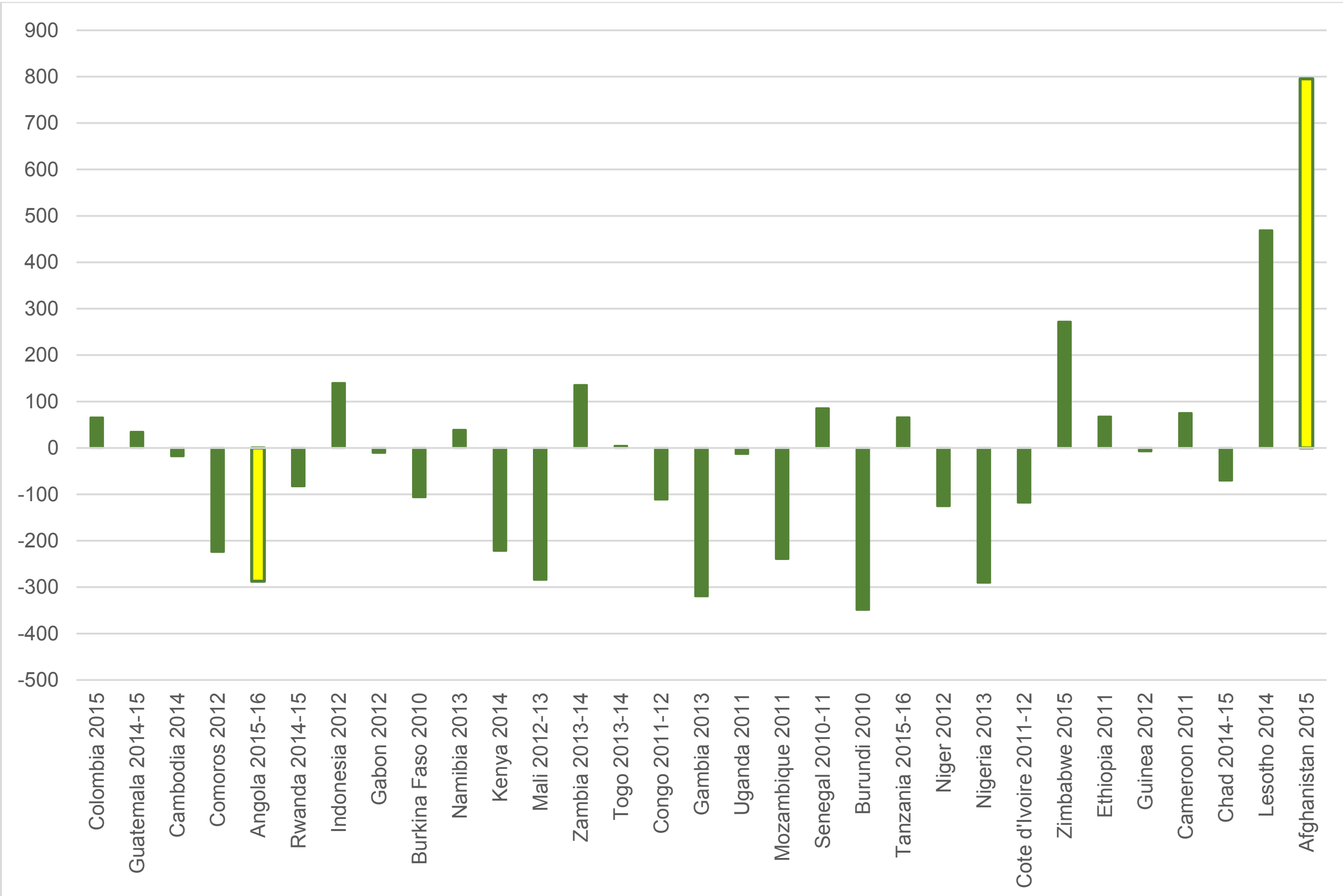


**Figure A.18. MMR versus Adult Female Mortality**

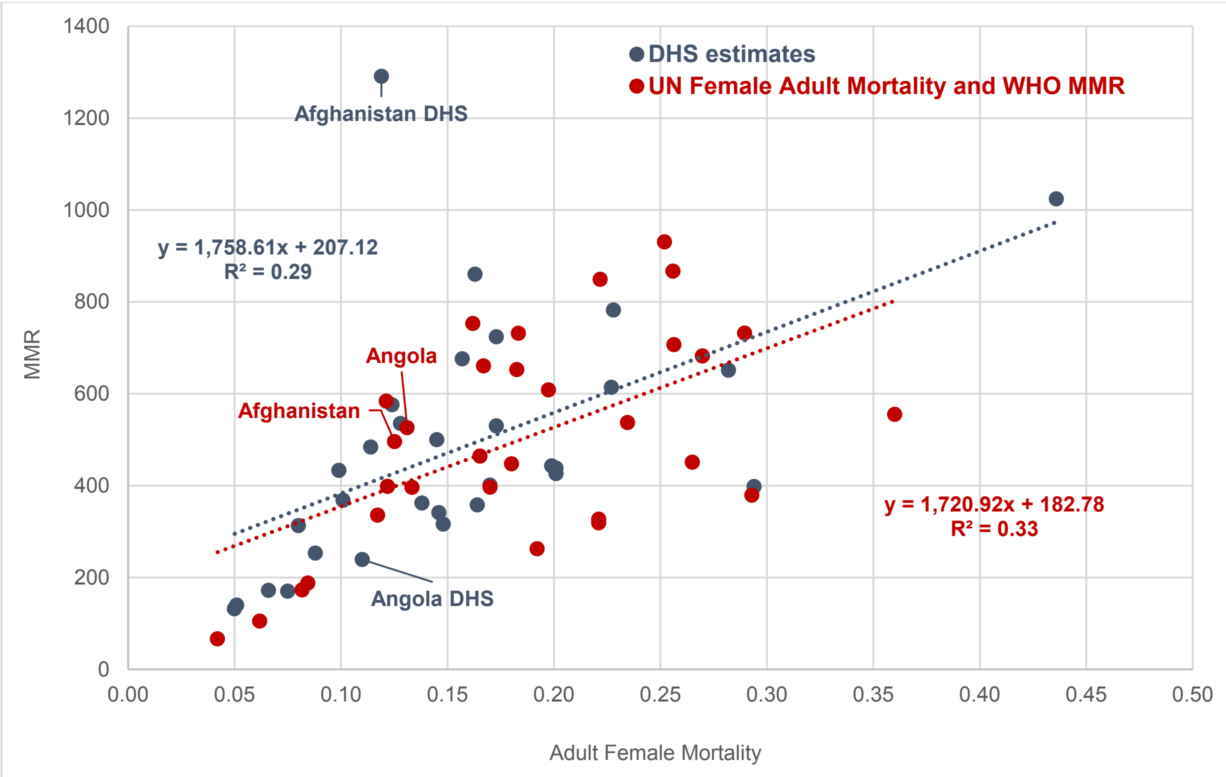



**Figure A.19. Difference of MMR versus difference of Adult Female Mortality**

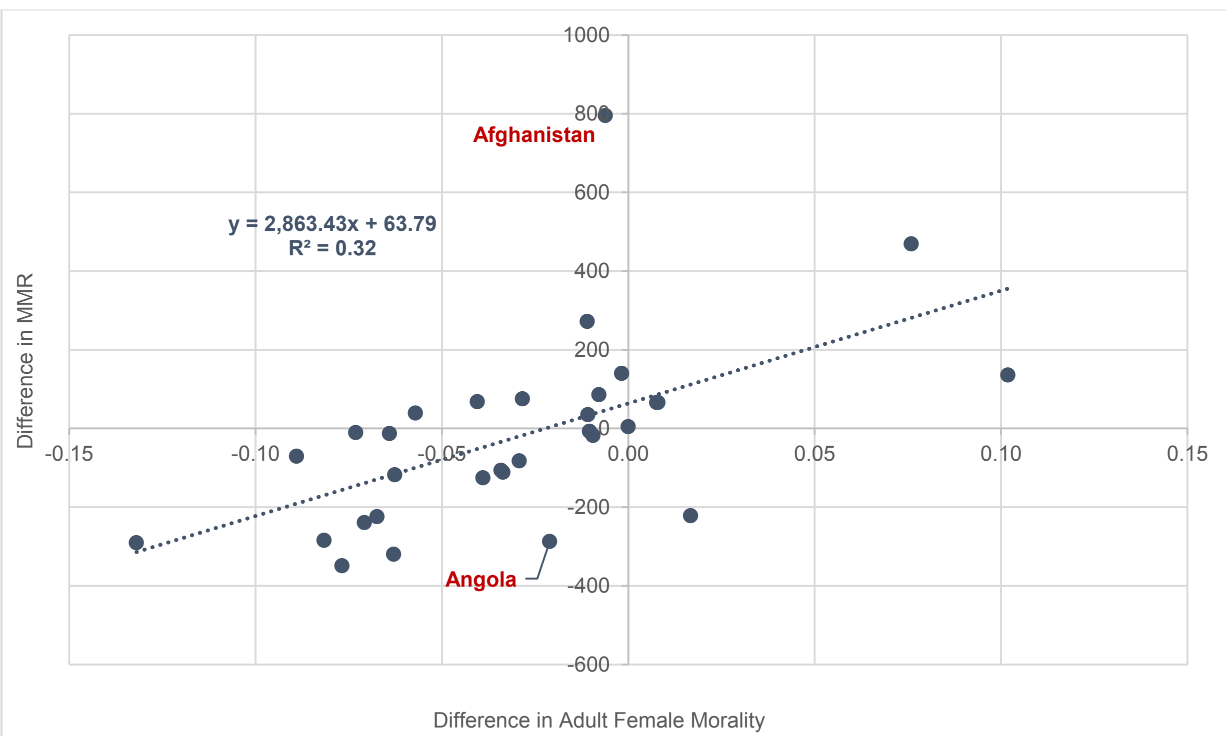


Figure A.20. U5MR versus MMR

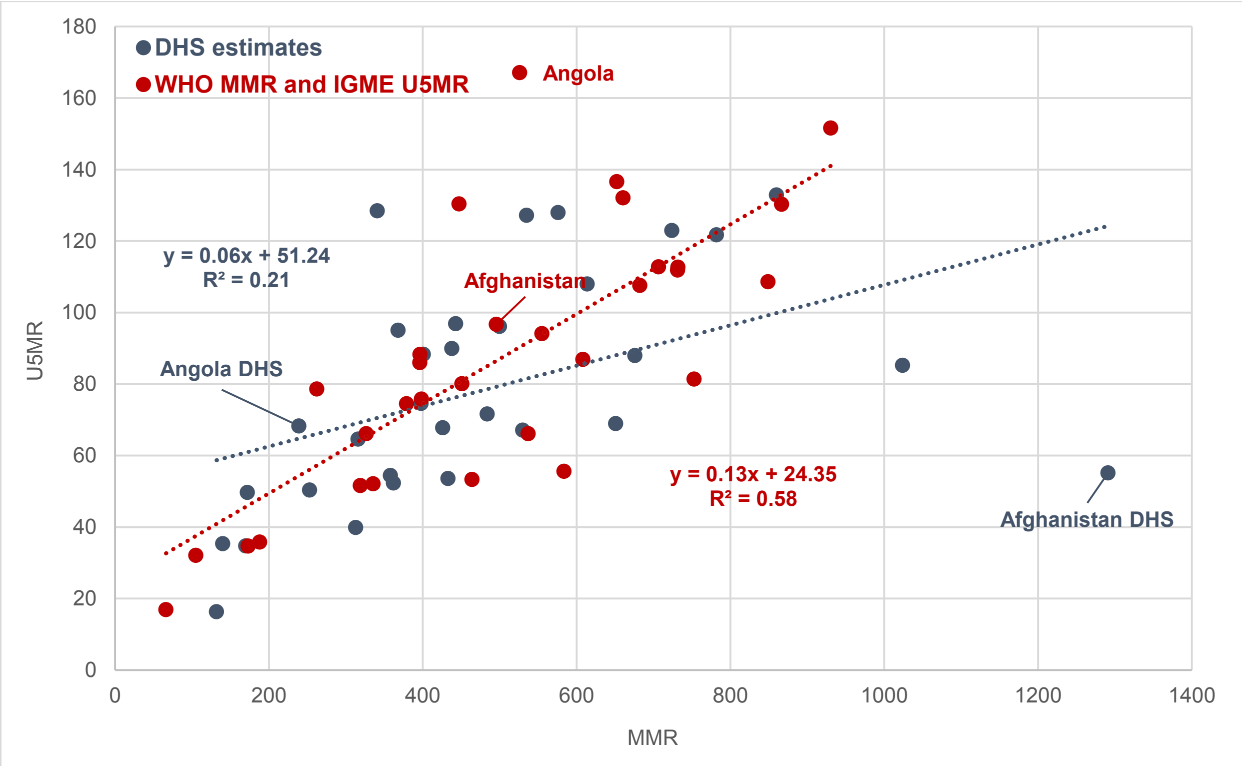



Figure A.21. Difference in U5MR versus difference in MMR

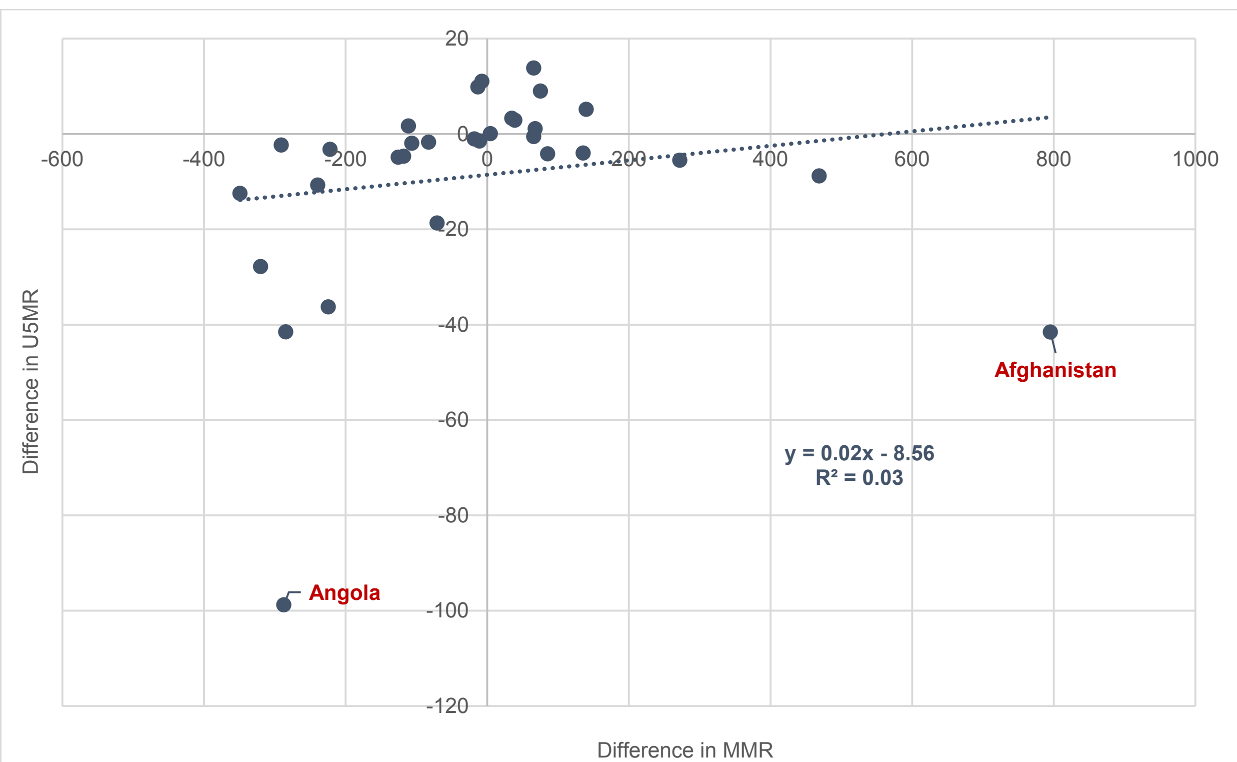


**Table A.2. TFR estimates from DHS and the UN Population Division**

| Survey | TFR [C.I.] DHS | TFR UN | Difference |
|---|---|---|---|
| Afghanistan 2015 | 5.29 [5.13,5.44] | 5.26 | 0.03 |
| Angola 2015-16 | 6.22 [6.02,6.42] | 5.95 | 0.27 |
| Armenia 2010 | 1.70 [1.58,1.84] | 1.72 | -0.02 |
| Bangladesh 2014 | 2.28 [2.19,2.38] | 2.22 | 0.06 |
| Benin 2011-12 | 4.90 [4.79,5.02] | 5.22 | -0.32 |
| Burkina Faso 2010 | 5.99 [5.85,6.13] | 6.08 | -0.09 |
| Burundi 2010 | 6.38 [6.20,6.57] | 6.50 | -0.12 |
| Cambodia 2014 | 2.72 [2.63,2.82] | 2.70 | 0.02 |
| Cameroon 2011 | 5.09 [4.94,5.24] | 5.25 | -0.16 |
| Chad 2014-15 | 6.45 [6.28,6.61] | 6.31 | 0.14 |
| Colombia 2015 | 1.97 [1.88,2.07] | 1.93 | 0.04 |
| Comoros 2012 | 4.32 [4.07,4.59] | 4.60 | -0.28 |
| Congo 2011-12 | 5.11 [4.93,5.30] | 4.86 | 0.25 |
| Cote d'Ivoire 2011-12 | 4.96 [4.73,5.19] | 5.14 | -0.18 |
| Dominican Republic 2013 | 2.48 [2.36,2.61] | 2.53 | -0.05 |
| DRC 2013-14 | 6.57 [6.37,6.76] | 6.40 | 0.17 |
| Egypt 2014 | 3.47 [3.39,3.54] | 3.38 | 0.09 |
| Ethiopia 2011 | 4.80 [4.55,5.07] | 5.26 | -0.46 |
| Gabon 2012 | 4.10 [3.89,4.33] | 4.00 | 0.10 |
| Gambia 2013 | 5.60 [5.40,5.82] | 5.62 | -0.02 |
| Ghana 2014 | 4.19 [4.00,4.40] | 4.18 | 0.01 |
| Guatemala 2014-15 | 3.13 [3.03,3.23] | 3.19 | -0.06 |
| Guinea 2012 | 5.10 [4.92,5.29] | 5.13 | -0.03 |
| Haiti 2012 | 3.53 [3.38,3.69] | 3.13 | 0.40 |
| Honduras 2011-12 | 2.94 [2.85,3.02] | 2.65 | 0.29 |
| Indonesia 2012 | 2.60 [2.53,2.67] | 2.45 | 0.15 |
| Jordan 2012 | 3.51 [3.38,3.64] | 3.60 | -0.09 |
| Kenya 2014 | 3.90 [3.79,4.02] | 4.10 | -0.20 |
| Kyrgyz Republic 2012 | 3.63 [3.46,3.82] | 3.12 | 0.52 |
| Lesotho 2014 | 3.26 [3.09,3.44] | 3.26 | 0.00 |
| Liberia 2013 | 4.73 [4.53,4.94] | 4.83 | -0.10 |
| Malawi 2015-16 | 4.43 [4.30,4.58] | 4.88 | -0.44 |
| Mali 2012-13 | 6.10 [5.92,6.28] | 6.35 | -0.25 |
| Mozambique 2011 | 5.92 [5.76,6.09] | 5.45 | 0.47 |
| Myanmar 2015-16 | 2.28 [2.16,2.41] | 2.30 | -0.02 |
| Namibia 2013 | 3.65 [3.50,3.80] | 3.60 | 0.05 |
| Nepal 2011 | 2.60 [2.42,2.80] | 2.96 | -0.36 |
| Niger 2012 | 7.64 [7.46,7.82] | 7.40 | 0.24 |
| Nigeria 2013 | 5.55 [5.41,5.69] | 5.74 | -0.19 |
| Pakistan 2012-13 | 3.83 [3.72,3.95] | 3.72 | 0.11 |
| Philippines 2013 | 3.04 [2.94,3.15] | 3.05 | -0.01 |
| Rwanda 2014-15 | 4.16 [4.04,4.29] | 4.20 | -0.04 |
| Senegal 2010-11 | 4.98 [4.82,5.16] | 5.10 | -0.12 |
| Sierra Leone 2013 | 4.91 [4.71,5.12] | 4.84 | 0.07 |
| Tajikistan 2012 | 3.76 [3.60,3.92] | 3.50 | 0.26 |
| Tanzania 2015-16 | 5.20 [5.01,5.40] | 5.24 | -0.04 |
| Togo 2013-14 | 4.78 [4.60,4.97] | 4.69 | 0.09 |
| Uganda 2011 | 6.20 [5.99,6.42] | 5.91 | 0.29 |
| Yemen 2013 | 4.43 [4.27,4.59] | 4.40 | 0.03 |
| Zambia 2013-14 | 5.26 [5.12,5.41] | 5.20 | 0.06 |
| Zimbabwe 2015 | 4.02 [3.88,4.17] | 4.00 | 0.02 |

Notes: Difference is DHS – UN estimate. UN estimate source (United Nations 2017).

**Table A.3. IMR and U5MR estimates from DHS and IGME**

| Survey | IMR [C.I.] DHS | IMR [U.I.][1] IGME | Difference IMR | U5MR [C.I.] DHS | U5MR [U.I.][1] IGME | Difference U5MR |
|---|---|---|---|---|---|---|
| Afghanistan 2015 | 44.8 [39.4,50.8] | 69.9 [58.7,83.5] | -25.1 | 55.1 [49.4,61.5] | 96.7 [79.1,118.2] | -41.6 |
| Angola 2015-16 | 44.3 [39.4,49.6] | 101.4 [68.8,149.2] | -57.2 | 68.3 [61.5,75.8] | 167.1 [107.7,252.2] | -98.8 |
| Armenia 2010 | 13.4 [7.8,22.9] | 17.9 [15.4,20.5] | -4.5 | 16.4 [10.3,26.1] | 20.0 [17.2,23.0] | -3.6 |
| Bangladesh 2014 | 38.1 [33.6,43.2] | 35.3 [32.7,38.0] | 2.8 | 46.1 [41.0,51.8] | 44.0 [40.3,47.8] | 2.1 |
| Benin 2011-12 | 41.8 [38.0,46.0] | 72.3 [63.8,80.3] | -30.5 | 70.1 [64.8,75.8] | 113.9 [98.7,128.2] | -43.8 |
| Burkina Faso 2010 | 65.1 [60.3,70.2] | 75.8 [71.2,80.8] | -10.7 | 128.4 [121.2,136.1] | 130.4 [117.9,144.3] | -2.0 |
| Burundi 2010 | 58.7 [52.5,65.6] | 69.3 [57.0,84.5] | -10.6 | 96.1 [88.2,104.6] | 108.6 [86.8,135.8] | -12.5 |
| Cambodia 2014 | 28.0 [22.8,34.2] | 30.7 [25.2,37.2] | -2.7 | 34.7 [28.8,41.8] | 35.8 [29.3,43.7] | -1.1 |
| Cameroon 2011 | 62.5 [57.0,68.4] | 70.8 [55.7,89.0] | -8.3 | 121.8 [114.3,129.6] | 112.8 [85.4,146.1] | 9.0 |
| Chad 2014-15 | 72.3 [66.7,78.2] | 90.2 [75.6,109.7] | -18.0 | 132.9 [125.4,140.7] | 151.6 [114.3,199.9] | -18.7 |
| Colombia 2015 | 13.9 [11.4,17.0] | 14.5 [10.7,19.9] | -0.6 | 16.3 [13.6,19.6] | 16.9 [12.4,23.4] | -0.6 |
| Comoros 2012 | 35.8 [28.1,45.6] | 63.1 [41.4,98.6] | -27.3 | 49.7 [39.5,62.3] | 86 [52.8,142.1] | -36.3 |
| Congo 2011-12 | 39.4 [33.5,46.2] | 45.3 [39.5,52.1] | -5.9 | 67.8 [59.9,76.6] | 66.1 [55.8,78.1] | 1.7 |
| Cote d'Ivoire 2011-12 | 67.9 [59.6,77.4] | 79.0 [71.3,87.7] | -11.1 | 108 [98.4,118.4] | 112.7 [100.4,126.6] | -4.7 |
| Dominican Republic 2013 | 27.2 [20.6,35.8] | 27.5 [23.6,32.3] | -0.3 | 30.9 [24.0,39.8] | 33.3 [28.1,39.8] | -2.4 |
| DRC 2013-14 | 58.3 [53.7,63.4] | 82.6 [70.2,94.8] | -24.3 | 104.2 [97.5,111.3] | 112.2 [91.4,135.4] | -8.0 |
| Egypt 2014 | 22.2 [19.4,25.5] | 23.4 [20.5,26.6] | -1.2 | 27.5 [24.3,31.0] | 27.9 [24.2,32.0] | -0.4 |
| Ethiopia 2011 | 59.2 [53.0,65.9] | 57.1 [50.4,64.4] | 2.1 | 88.0 [79.8,97.0] | 86.9 [75.0,99.8] | 1.1 |
| Gabon 2012 | 42.5 [35.7,50.5] | 44.2 [38.5,50.7] | -1.7 | 64.6 [56.5,73.8] | 66.1 [55.6,78.1] | -1.5 |
| Gambia 2013 | 34.3 [29.1,40.5] | 51.7 [45.9,59.2] | -17.4 | 53.6 [45.8,62.6] | 81.4 [62.0,105.6] | -27.8 |
| Ghana 2014 | 41.2 [35.1,48.4] | 47.0 [41.0,53.8] | -5.8 | 59.9 [52.6,68.0] | 69.2 [58.5,81.1] | -9.4 |
| Guatemala 2014-15 | 27.8 [24.3,31.8] | 26.6 [19.9,35.6] | 1.2 | 35.4 [31.4,39.8] | 32.1 [23.4,44.5] | 3.3 |
| Guinea 2012 | 66.8 [59.5,75.0] | 71.2 [62.6,81.2] | -4.4 | 122.9 [112.8,133.8] | 111.9 [96.6,129.8] | 11.0 |
| Haiti 2012 | 59.2 [52.7,66.6] | 60.0 [54.5,66.6] | -0.8 | 88.4 [79.6,98.2] | 81.2 [72.5,91.6] | 7.2 |
| Honduras 2011-12 | 23.6 [20.3,27.4] | 21.4 [18.9,24.3] | 2.2 | 29.0 [25.5,33.0] | 25.4 [22.2,29.1] | 3.6 |
| Indonesia 2012 | 31.5 [27.7,35.8] | 28.5 [26.3,30.8] | 3.0 | 39.9 [35.6,44.6] | 34.7 [31.8,37.7] | 5.2 |
| Jordan 2012 | 17.2 [13.0,22.7] | 17.8 [15.1,21.0] | -0.6 | 21.0 [16.5,26.7] | 20.8 [17.5,24.8] | 0.2 |
| Kenya 2014 | 38.7 [35.0,42.8] | 39.2 [33.5,45.7] | -0.5 | 52.3 [48.2,56.8] | 55.6 [45.8,67.2] | -3.3 |
| Kyrgyz Republic 2012 | 26.9 [21.1,34.1] | 26.5 [25.8,27.1] | 0.4 | 31.0 [24.9,38.4] | 30.0 [29.2,30.8] | 1.0 |
| Lesotho 2014 | 59.5 [50.3,70.3] | 72.3 [62.2,83.9] | -12.8 | 85.3 [74.3,97.7] | 94.1 [79.4,111.3] | -8.8 |
| Liberia 2013 | 53.8 [47.0,61.4] | 65.2 [58.6,72.8] | -11.4 | 93.8 [84.5,104.0] | 89.3 [78.9,101.2] | 4.5 |
| Malawi 2015-16 | 41.7 [37.8,45.9] | 47.3 [39.9,58.8] | -5.6 | 63.8 [58.9,69.0] | 71.3 [57.8,92.4] | -7.5 |
| Mali 2012-13 | 56.0 [49.8,62.9] | 82.9 [71.0,97.7] | -26.9 | 95.1 [86.6,104.3] | 136.6 [105.3,174.3] | -41.5 |
| Mozambique 2011 | 64.1 [58.3,70.4] | 74.3 [67.8,81.4] | -10.2 | 96.9 [89.8,104.4] | 107.6 [97.0,119.3] | -10.7 |
| Myanmar 2015-16 | 40.3 [33.2,48.9] | 41.9 [34.2,50.6] | -1.6 | 50.0 [41.9,59.6] | 53.5 [42.5,66.5] | -3.5 |
| Namibia 2013 | 38.9 [32.7,46.1] | 36.0 [30.5,42.9] | 2.9 | 54.5 [47.1,62.9] | 51.6 [42.3,64.2] | 2.9 |
| Nepal 2011 | 45.9 [39.2,53.7] | 40.0 [36.2,44.1] | 5.9 | 54.4 [47.1,62.7] | 50.7 [45.2,56.7] | 3.7 |
| Niger 2012 | 50.6 [45.4,56.5] | 68.9 [64.1,74.3] | -18.3 | 127.2 [119.3,135.5] | 132.1 [117.3,148.8] | -4.9 |
| Nigeria 2013 | 68.5 [64.6,72.7] | 81.5 [73.4,90.3] | -13.0 | 128.0 [121.7,134.5] | 130.3 [115.8,146.4] | -2.3 |
| Pakistan 2012-13 | 73.6 [66.8,81.0] | 73.5 [66.9,81.4] | 0.1 | 88.9 [80.8,97.8] | 91.8 [82.6,103.1] | -2.9 |
| Philippines 2013 | 22.7 [19.3,26.6] | 24.4 [21.1,27.9] | -1.7 | 31.1 [27.0,35.8] | 31.2 [26.5,36.5] | -0.1 |
| Rwanda 2014-15 | 32.3 [28.2,37.0] | 37.1 [32.1,42.9] | -4.8 | 50.4 [45.0,56.4] | 52.1 [43.6,62.3] | -1.7 |
| Senegal 2010-11 | 46.7 [42.3,51.6] | 50.0 [47.8,52.4] | -3.3 | 71.6 [65.6,78.1] | 75.8 [68.5,83.7] | -4.2 |
| Sierra Leone 2013 | 92.3 [85.6,99.6] | 102.3 [94.5,110.9] | -10.0 | 155.8 [146.7,165.3] | 150.6 [134.8,168.3] | 5.2 |
| Tajikistan 2012 | 34.5 [28.9,41.2] | 44.7 [36.4,54.6] | -10.2 | 155.8 [146.7,165.3] | 150.6 [134.8,168.3] | 5.2 |
| Tanzania 2015-16 | 43.2 [38.5,48.5] | 37.6 [29.6,47.6] | 5.6 | 43.4 [36.8,51.1] | 52.6 [42.2,65.8] | -9.2 |
| Togo 2013-14 | 48.5 [42.6,55.2] | 57.9 [52.2,64.0] | -9.4 | 67.1 [60.7,74.1] | 53.3 [39.8,71.2] | 13.8 |
| Uganda 2011 | 53.8 [48.2,60.0] | 52.3 [48.0,57.1] | 1.5 | 88.3 [79.8,97.7] | 88.3 [78.3,99.1] | 0.0 |
| Yemen 2013 | 43.2 [38.8,48.1] | 40.3 [35.1,46.4] | 2.9 | 90 [81.9,98.8] | 80.1 [72.2,88.9] | 9.9 |
| Zambia 2013-14 | 44.6 [40.2,49.6] | 51.1 [46.9,56.1] | -6.5 | 52.7 [47.7,58.3] | 51.1 [43.7,60.1] | 1.6 |
| Zimbabwe 2015 | 50.1 [44.4,56.7] | 48.8 [40.4,59.9] | 1.3 | 69 [61.1,77.7] | 74.5 [59.1,95.1] | -5.5 |

[1]90% uncertainty intervals. Differences are the DHS – IGME estimates. IGME source (UNICEF 2017).

**Table A.4. Adult mortality probabilities (15q35*1000) from DHS and the UN Population Division**

| Survey | Female Adult Mortality DHS | Female Adult Mortality UN | Difference Female | Male Adult Mortality DHS | Male Adult Mortality UN | Difference Male |
|---|---|---|---|---|---|---|
| Afghanistan 2015 | 119 | 125 | -6 | 84 | 140 | -56 |
| Angola 2015-16 | 110 | 131 | -21 | 182 | 180 | 2 |
| Burkina Faso 2010 | 146 | 180 | -34 | 145 | 186 | -41 |
| Burundi 2010 | 145 | 222 | -77 | 179 | 250 | -71 |
| Cambodia 2014 | 75 | 84 | -9 | 129 | 126 | 3 |
| Cameroon 2011 | 228 | 256 | -28 | 232 | 262 | -30 |
| Chad 2014-15 | 163 | 252 | -89 | 189 | 273 | -84 |
| Colombia 2015 | 50 | 42 | 8 | 105 | 126 | -21 |
| Comoros 2012 | 66 | 133 | -67 | 66 | 155 | -89 |
| Congo 2011-12 | 201 | 235 | -34 | 173 | 239 | -66 |
| Cote d'Ivoire 2011-12 | 227 | 290 | -63 | 205 | 303 | -98 |
| Ethiopia 2011 | 157 | 197 | -40 | 181 | 221 | -40 |
| Gabon 2012 | 148 | 221 | -73 | 171 | 200 | -29 |
| Gambia 2013 | 99 | 162 | -63 | 102 | 185 | -83 |
| Guatemala 2014-15 | 51 | 62 | -11 | 115 | 144 | -29 |
| Guinea 2012 | 173 | 183 | -10 | 173 | 197 | -24 |
| Indonesia 2012 | 80 | 82 | -2 | 97 | 105 | -8 |
| Kenya 2014 | 138 | 121 | 17 | 183 | 179 | 4 |
| Lesotho 2014 | 436 | 360 | 76 | 476 | 429 | 47 |
| Malawi 2015-16 | 184 | 193 | -9 | 218 | 252 | -34 |
| Mali 2012-13 | 101 | 183 | -82 | 105 | 188 | -83 |
| Mozambique 2011 | 199 | 270 | -71 | 241 | 309 | -68 |
| Myanmar 2015-16 | 72 | 89 | -17 | 163 | 118 | 45 |
| Namibia 2013 | 164 | 221 | -57 | 244 | 265 | -21 |
| Niger 2012 | 128 | 167 | -39 | 136 | 176 | -40 |
| Nigeria 2013 | 124 | 256 | -132 | 123 | 263 | -140 |
| Rwanda 2014-15 | 88 | 117 | -29 | 124 | 151 | -27 |
| Senegal 2010-11 | 114 | 122 | -8 | 115 | 143 | -28 |
| Tanzania 2015-16 | 173 | 165 | 8 | 181 | 217 | -36 |
| Togo 2013-14 | 170 | 170 | 0 | 176 | 176 | 0 |
| Uganda 2011 | 201 | 265 | -64 | 252 | 299 | -47 |
| Zambia 2013-14 | 294 | 192 | 102 | 330 | 246 | 84 |
| Zimbabwe 2015 | 282 | 293 | -11 | 300 | 315 | -15 |

Notes: Difference is DHS – UN estimate. UN source (United Nations 2017).

**Table A.5. MMR estimates from DHS and WHO**

| Survey | MMR [C.I.] DHS | MMR [U.I.][1] WHO | Difference |
|---|---|---|---|
| Afghanistan 2015 | 1291 [1071,1512] | 496 [325,752] | 795 |
| Angola 2015-16 | 239 [164,313] | 526 [251,1048] | -287 |
| Burkina Faso 2010 | 341 [275,406] | 447 [350,563] | -106 |
| Burundi 2010 | 500 [376,624] | 849 [622,1145] | -349 |
| Cambodia 2014 | 170 [95,246] | 188 [148,240] | -18 |
| Cameroon 2011 | 782 [647,916] | 707 [558,924] | 75 |
| Chad 2014-15 | 860 [728,993] | 931 [635,1382] | -71 |
| Colombia 2015 | 132 [66,198] | 66 [59,79] | 66 |
| Comoros 2012 | 172 [60,284] | 396 [259,610] | -224 |
| Congo 2011-12 | 426 [274,579] | 537 [398,718] | -111 |
| Cote d'Ivoire 2011-12 | 614 [445,783] | 732 [551,952] | -118 |
| Ethiopia 2011 | 676 [541,810] | 608 [470,824] | 68 |
| Gabon 2012 | 316 [178,454] | 327 [236,451] | -11 |
| Gambia 2013 | 433 [299,567] | 753 [542,1030] | -320 |
| Guatemala 2014-15 | 140 [94,186] | 105 [99,112] | 35 |
| Guinea 2012 | 724 [531,916] | 731 [574,935] | -7 |
| Indonesia 2012 | 313 [220,406] | 173 [135,223] | 140 |
| Kenya 2014 | 362 [254,471] | 584 [423,817] | -222 |
| Lesotho 2014 | 1024 [731,1318] | 555 [366,935] | 469 |
| Mali 2012-13 | 368 [259,478] | 652 [532,820] | -284 |
| Mozambique 2011 | 443 [328,559] | 682 [539,864] | -239 |
| Namibia 2013 | 358 [222,495] | 319 [230,437] | 39 |
| Niger 2012 | 535 [425,645] | 661 [531,830] | -126 |
| Nigeria 2013 | 576 [500,652] | 867 [673,1128] | -291 |
| Rwanda 2014-15 | 253 [181,326] | 336 [256,430] | -83 |
| Senegal 2010-11 | 484 [337,631] | 398 [286,551] | 86 |
| Tanzania 2015-16 | 530 [405,655] | 464 [342,638] | 66 |
| Togo 2013-14 | 401 [290,512] | 396 [285,540] | 5 |
| Uganda 2011 | 438 [368,507] | 451 [354,583] | -13 |
| Zambia 2013-14 | 398 [323,474] | 262 [204,330] | 136 |
| Zimbabwe 2015 | 651 [473,829] | 379 [324,454] | 272 |

Notes: [1]80% uncertainty intervals. Difference is DHS – WHO estimate. WHO source (WHO 2015).