Three simple tools to evaluate the quality of fertility estimates from birth histories

Application to African Demographic and Health surveys

Bruno Schoumaker, Université catholique de Louvain

Asbtract for the 2015 UAPS Population Conference

Introduction

Fertility estimates are central to designing and evaluating population policy, establishing population projections, and documenting the dynamics of population. Such objectives require good quality data. Despite the increasing availability of fertility data in sub-Saharan Africa, especially thanks to the Demographic and Health Surveys program, estimates remain affected by data quality problems. Underreporting of births, displacements of births and issues with sample implementation may lead to more or less severe biases in fertility levels and trends (Schoumaker, 2014a). Recent discussion on stalls in fertility transitions in sub-Saharan Africa (Schoumaker, 2009), or more ancient debates about the genuineness of early declines (Blacker, 1994) indicate that issues of measurements may have important policy implications.

However, detecting and measuring measurement errors in fertility estimates is not straightforward. While some routine indicators are used to detect typical data quality issues in surveys (such as heaping), tools for evaluating the quality of fertility estimates from birth histories are few, and - to varying extent – unsatisfactory. For instance, underreporting of recent births is typically detected by examining sex ratios at birth or ratios of neonatal to postneonatal mortality; these indicators are limited to detecting selective underreporting (Curtis, 1995). The application of methods recently developed in DHS methodological reports are not straightforward (Pullum & Becker, 2014; Schoumaker, 2014a), and these methods may not be easily adopted by survey users.

In this paper, I present three simple tools that can be implemented easily to identify potential data quality problems in fertility estimates. The tools necessitate at least two surveys. This is the case in most countries. The three tools rely on a simple idea: good quality data provide consistent estimates across surveys or across methods; in consequence, inconsistencies across surveys or across methods reflecting data quality issues. The three tools are based on this same idea, but use different methods, and as a result are sensitive to different types of data quality problems. Used together, these tools enable researchers to obtain a quick diagnostic of the quality of fertility estimates. In summary, they can be used to detect data quality issues and to some extent to point to specific types of errors. Stata commands are provided and allow users to apply these tools in a user-friendly way.

Data

Data come from the Demographic and Heath Surveys program, which is the main source of data on fertility in sub-Saharan Africa. Fertility data are collected through full birth histories among women aged 15-49. A wide variety of other topics are also covered by the DHS. Of particular relevance is and the lengthy module on children's health, that focuses on children born in the last 5 years. This is one of the causes of measurement errors of recent fertility, as interviewers may be tempted to displace births backward, or to omit births altogether in order to avoid administering the health module. I also use simulated individual birth histories produced with the SOCSIM Software. These simulated histories are used to test the tools in a controlled environment, either without measurement errors or with known measurement errors.

Preliminary results presented here are based on data from selected sub-Saharan African countries; analyses will be performed for all the DHS, as well as some MICS surveys.

Methods

Published fertility rates in Demographic and Health surveys are usually computed for the three years preceding the survey among women aged 15-49 (orange area, Figure 1). Fertility trends are typically measured by comparing these three-year estimates in successive surveys.

Figure 1 : Lexis diagram illustrating the computation of recent fertility in DHS

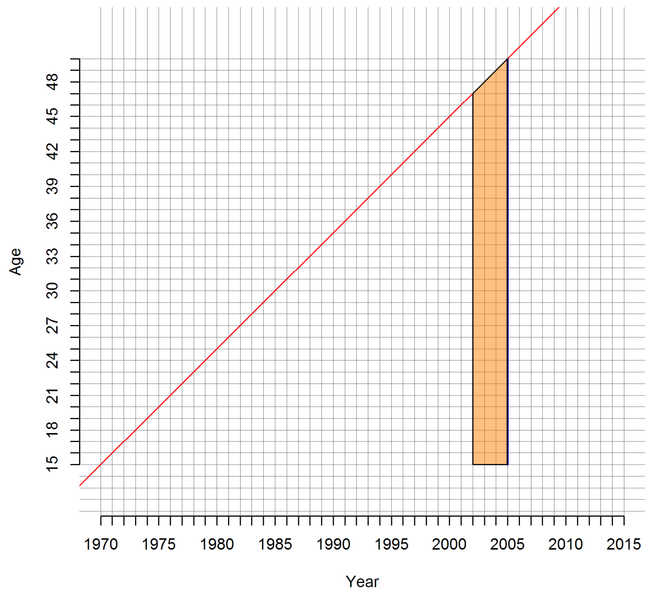
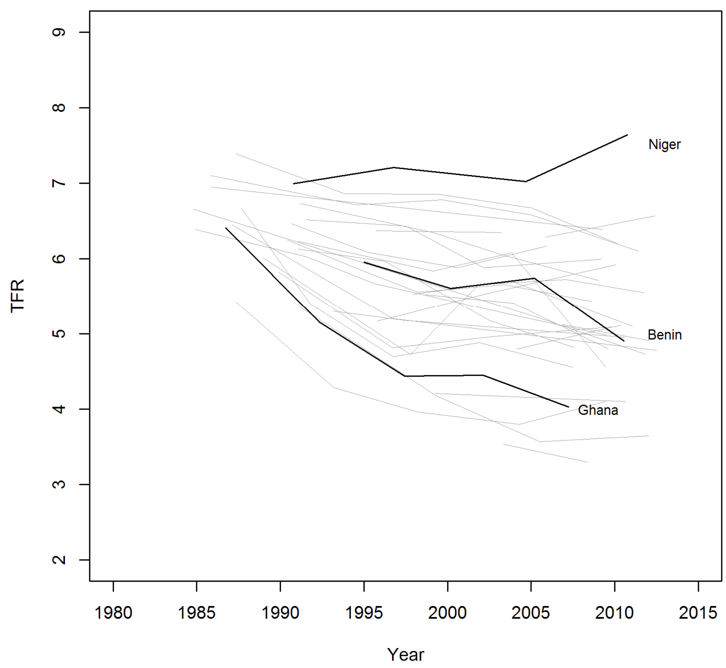


Figure 2 shows these fertility trends using all African Demographic and Health Surveys. A few countries are highlighted and will be discussed in preliminary results. (1) Niger, with a high and slightly increasing fertility; (2) Benin, with an apparent rapid decrease in fertility; (3) Ghana, that where a stall is visible in the early 2000s..

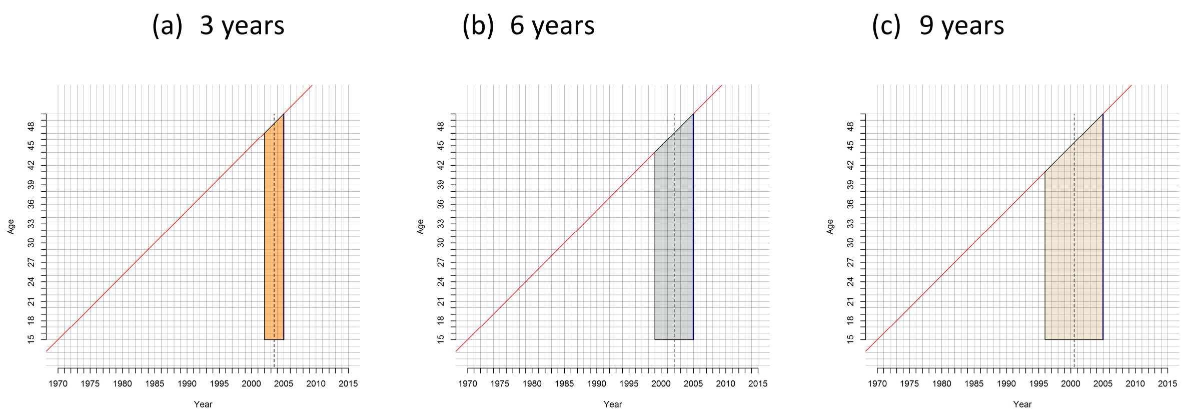
Figure 2 : Fertility trends bases on three-year estimates in successive DHS



Tool 1: changing the reference period

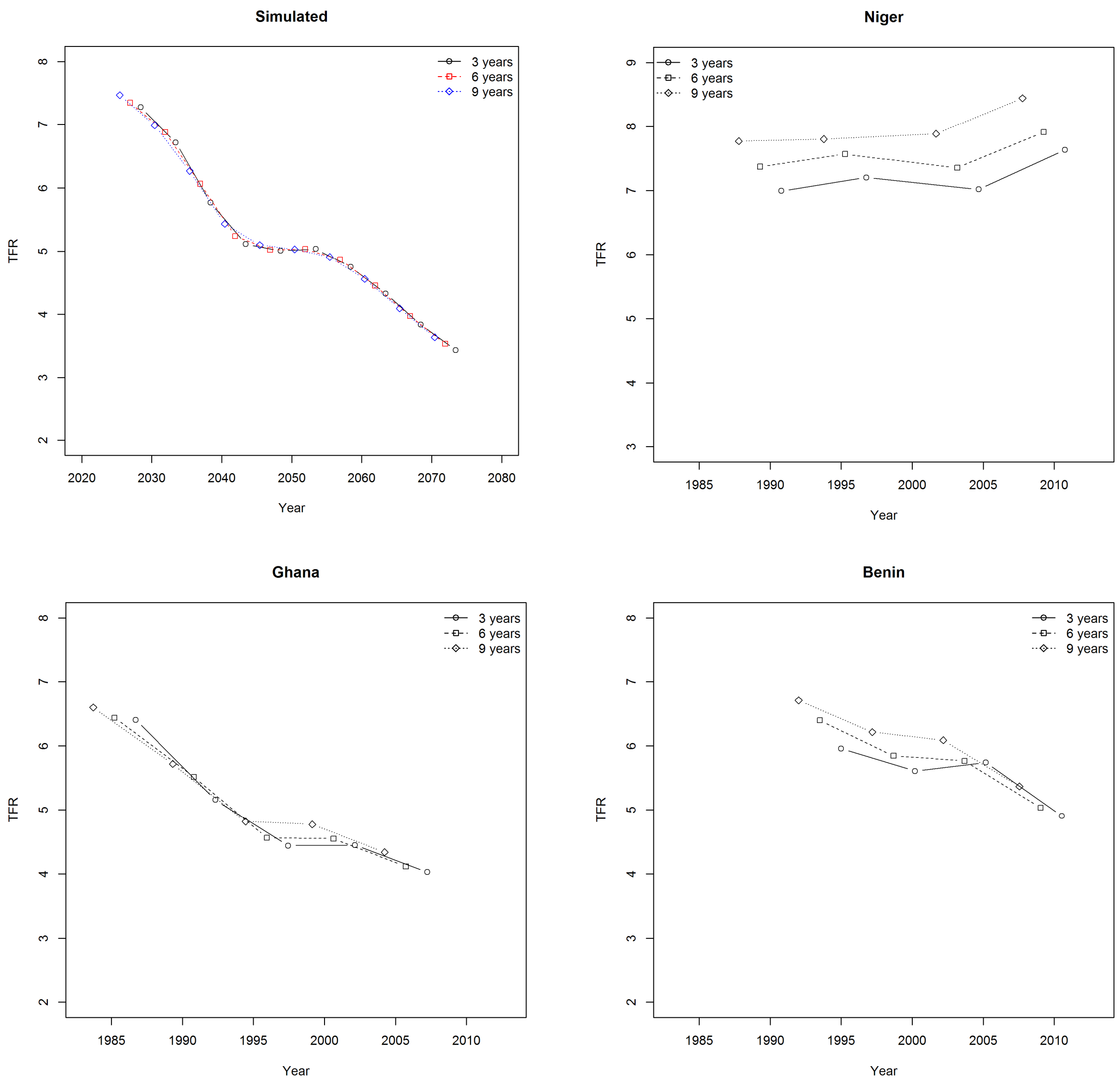
A first and very simple tool consists in comparing fertility estimates computed with different reference periods. Published TFRs are computed for the three years preceding the surveys. Since full birth histories are available, fertility rates can be computed over periods of other lengths, for instance 6 years or 9 years, as illustrated on Figure 3.

Figure 3 : Lexis diagram illustrating the computation of recent fertility with different reference periods



The tool is applied to good quality simulated birth histories to show that – with good quality data - *fertility trends are insensitive to the length of the reference period* (Figure 4). In other words, fertility levels and trends should be similar regardless of the reference period.

Figure 4: comparison of fertility trends with different reference period, simulated birth histories and real birth histories from three countries

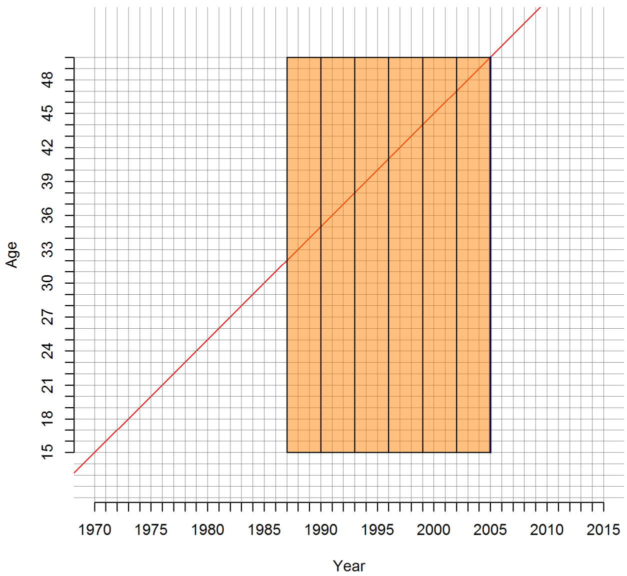


Application of this tool to real data shows that discrepancies across reference periods can be quite severe (Figure 4). For instance, changing the reference period leads to differences in fertility levels of one child in Niger, and more than half a child in early surveys in Benin. Ghana's example also shows data quality issues in the late 1990s/early 2000s, the period were the stall was observed. In Benin, discrepancies are also found in the first three surveys. This tool is thus a very simple way to show that there are data quality issues, and it easily conveys a sense of the magnitude of these problems. It does not show in a straightforward way the reasons for these discrepancies. Simulations can be used to show that such discrepancies results either from underreporting of recent births (the case in Niger), or from differences in sample implementation (a possible issue in Ghana). This will be developed in the full paper. The second and third tools provide additional insights into the causes of these discrepancies, and also allow detecting other issues.

Tool 2: comparing reconstructed fertility trends

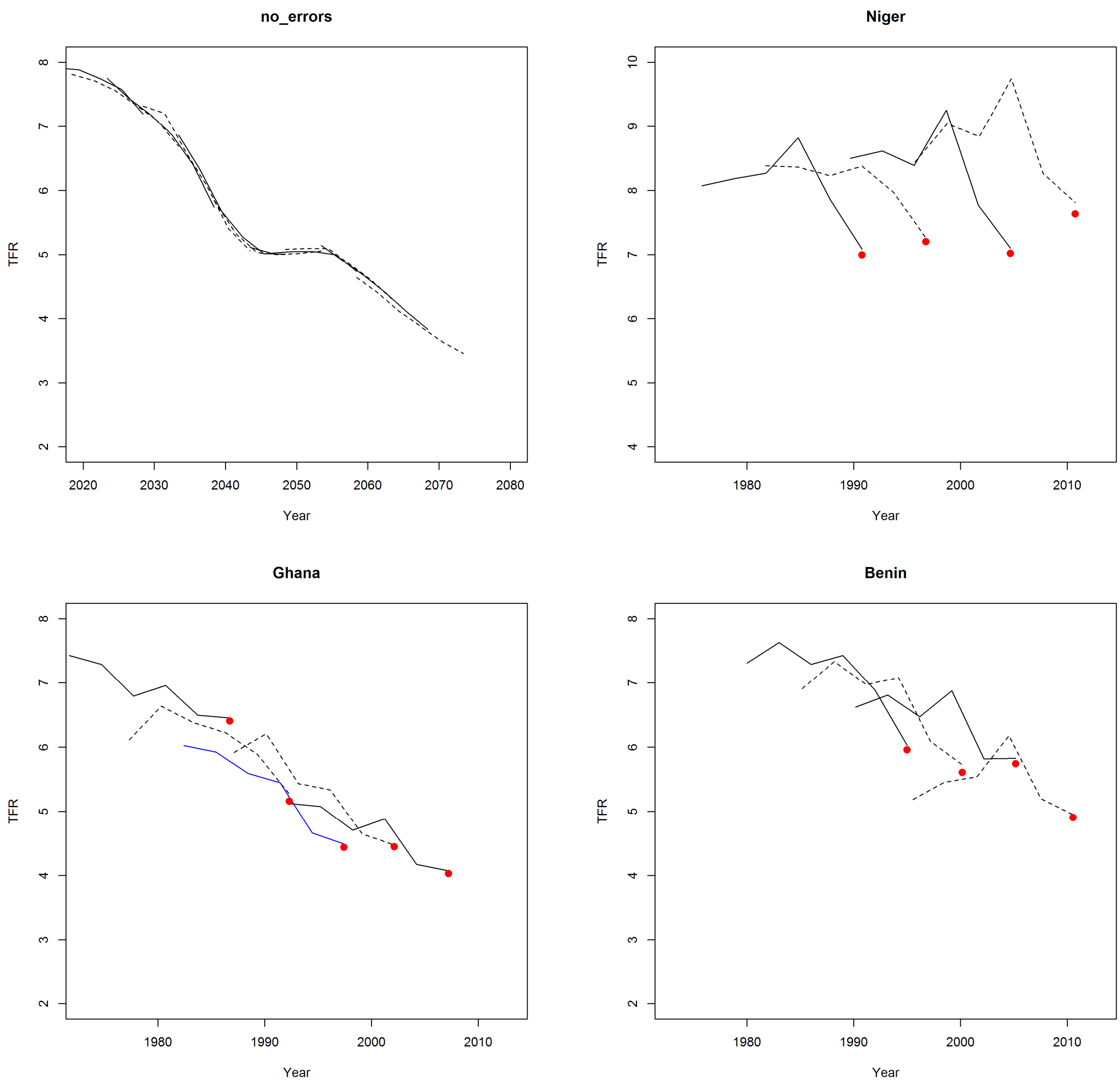
The second tool consists in reconstructing fertility trends by 3-year periods over the 18 years preceding each survey, and comparing trends across surveys. This is illustrated on Figure 5, where each rectangle corresponds to a 3-year period for which the total fertility rates are computed. Although the data is truncated, simulations show that the TFR (15-49) can be reconstructed in a reliable way over 15-20 years (Schoumaker, 2013). Fertility is reconstructed using Poisson regression. The method consists of creating a table of births and exposure by 3-year periods and by 5-year age groups of women from the birth history data. A Poisson model is then fitted with the number of births as the dependent variable, exposure as an offset, and two independent variables: 1) age measured with dummy variables for 5-year age groups, and 2) periods measured with dummy variables. The total fertility rate between age 15 and age 49 can be reconstructed for each period from the regression coefficients. The method for reconstructing fertility in this way is presented in detail in Schoumaker (2013), and a Stata module implementing the method is available.

Figure 5 : Lexis diagram illustrating the computation of fertility over the last 18 years by 3-year periods



Simulations are first used to show that good quality data leads to highly consistent reconstructed fertility estimates across surveys (Figure 6). Comparing reconstructing trends in Niger, Ghana and Benin confirm the discrepancies found with tool 1, indicate possible sources of discrepancies, and also point to issues not detected with the first tool, but that will be confirmed with the third tool. In Niger, recent estimates (illustrated with red dots) are much lower than retrospective estimates for the same periods from later surveys. The sudden drop in fertility in the first few years before the survey is consistent with underreporting of recent births, possibly due to the length health module. Both this tool and the first one suggest recent fertility is severely underestimated in Niger. A similar – but less pronounced - pattern is found in Benin. This tool also shows another feature of Benin's data: retrospective estimates in the latest survey appear lower than estimates from previous surveys. A possible cause of this is a difference in sample composition, with low fertility women being oversampled. The same type of issue is found in Ghana, where estimates in the third survey seem to be significantly lower than estimates from the other surveys.

Figure 6: comparison of fertility trends with different reference period, simulated birth histories and real birth histories from three countries



Tool 3: comparing direct and indirect estimates

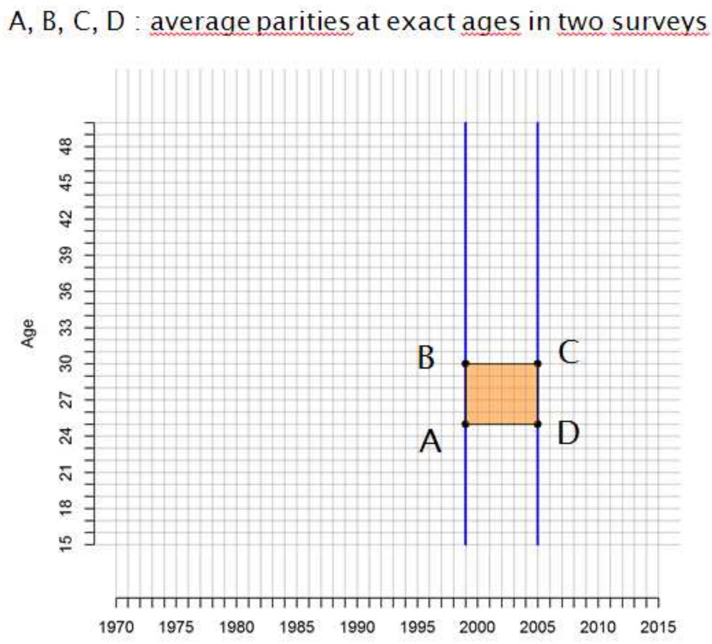
The third tool relies on the comparison of direct and indirect estimates of fertility. Once again, the starting point is that direct and indirect estimates should provide consistent estimates with good quality data. Differences across methods provide evidence of data quality issues.

The method we use consists in estimating of period age-specific fertility rates from the increment of cohort parities (Zlotnik & Hill, 1981). The central idea of that method is that the change in the parity in a cohort of women between two dates reflects the fertility rate during the time interval. The crisscross method (Schmertmann, 2002) can be viewed as a generalization of the earlier methods relying on cohort parity increments. With the crisscross method, the fertility rate (*Rate*) between two exact ages (x and x+n) over a period of any length t, is obtained in the following way (Eq. 1, Figure 7).

$$Rate = (C - A) \times \left(\frac{1}{2 \times n} + \frac{1}{2 \times t}\right) - (B - D) \times \left(\frac{1}{2 \times n} - \frac{1}{2 \times t}\right) \quad \text{(Eq.1)}$$

Where A, B, C and D are the mean number of children ever born (parities) at exact ages and dates defined by the corners of the Lexis diagram, *t* is the time interval between the two surveys, and *n* is the width of the age group (Figure 7).

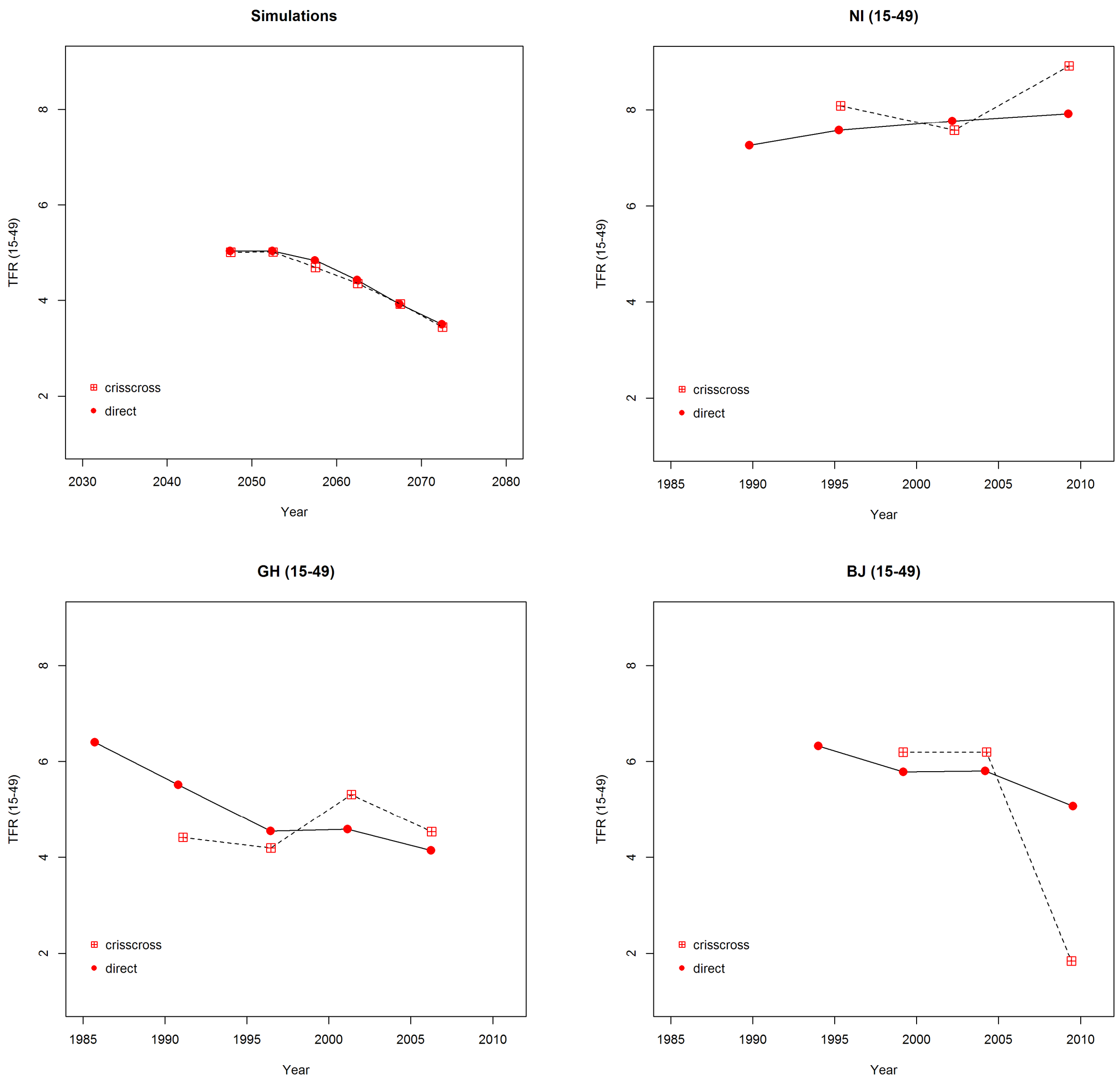
Figure 7: Lexis diagram illustrating the computation of fertility rates with the crisscross method



The method we use consists in comparing the indirect estimates, obtained from two surveys, with direct estimates (from the most recent survey), as done by Blacker (1994). As discussed by previous authors, the crisscross method tends to magnify data quality problems. Zlotnik and Hill (1981, p. 106) found that "any change in error from one survey to the another, will be exaggerated in the synthetic data set", and this should not be "regarded entirely as a vice, since their sensitivity to error makes the technique proposed very useful in detecting it". The method is especially sensitive to differences in sample composition. This is a key point of this method.

Simulations show again that indirect and direct estimates match quite well with good quality data (Figure 8). And once again, comparisons with real data indicate potentially serious data quality issues. Benin is an interesting case. The most recent estimate with the crisscross method is much lower than direct estimates and is clearly implausible. Such a pattern is typical of underestimation of parities in the second survey compared to the first survey (Schoumaker, 2014b), and suggests either omissions of births in the last survey, or oversampling of low parity women. The second interpretation is consistent with the result of tool 2, and it is fair to say that fertility was underestimated in the most recent survey because of sampling issues. As a result the recent decline in Benin is overestimated. The Ghana situation, with the dotted line crossing the solid line between the late 1990s and early 2000s also suggests that parities were underestimated in the survey in the late 1990s, possibility because of sampling issues. The stall in Ghana may be spurious. The third tool is less useful for Niger, but tends to confirm that recent births were omitted in Niger. The indirect estimates tend to be higher than direct estimates which is found when omission of recent births is constant across surveys (Schoumaker, 2014b).

Figure 8: comparison of fertility trends with different reference period, simulated birth histories and real birth histories from three countries



Further analyses

These three tools can be applied fairly easily with Stata. A Stata commands prepared by the author (tfr2) is available for tool 1 and tool 2 (Schoumaker, 2013). Another Stata command will be provided for the third tool.

Guidelines will also be provided to interpret these results will in a more systematic way. The basic idea consists in providing users simple guidelines to interpret jointly results from the three methods, and reach a quick assessment of the quality of fertility estimates and of the possible source of errors. The guidelines will be presented as a table with a limited number of entries describing results of the three tools, and a corresponding diagnostic of the probable source of errors. The preparation of this table will rely on simulating a large number of birth histories with known data quality problems to 'map' the links between errors and the results of tools 1, 2 and 3 in this controlled environment (Schoumaker, 2014b). Application to a series of surveys from sub-Saharan Africa will be used to illustrate the utility of the method, and show that estimates of fertility levels and trends should be interpreted with caution in sub-Saharan Africa.

Bibliography

Blacker, J. (1994). Some Thoughts on the Evidence of Fertility Decline in Eastern and Southern Africa. *Population and Development Review*, *20*(1), 200–205. http://doi.org/10.2307/2137639

Curtis, S. (1995). *Assessment of the Quality of Data Used for Direct Estimation of Infant and Child Mortality in DHS-II Surveys* (DHS Occasional Papers No. 3). Calverton: Macro International.

Pullum, T., & Becker, S. (2014). *Evidence of Omission and Displacement in DHS Birth Histories.* (DHS Methodological Report No. 11). Calve: ICF International.

Schmertmann, C. P. (2002). A Simple Method for Estimating Age-Specific Rates from Sequential Cross Sections. *Demography*, *39*(2), 287–310. http://doi.org/10.2307/3088340

Schoumaker, B. (2009). *Stalls in Fertility Transitions in Sub- Saharan Africa: Real or Spurious?* (Document de travail SPED No. 30). Louvain-la-Neuve: Université catholique de Louvain.

Schoumaker, B. (2013). A Stata module for computing fertility rates and TFRs from birth histories: tfr2. *Demographic Research*, *28*(38), 1093–1144. http://doi.org/10.4054/DemRes.2013.28.38

Schoumaker, B. (2014a). *Quality and Consistency of DHS Fertility Estimates, 1990-2012.* (DHS Methodological Report No. 12) (p. 106). Rockville: ICF International. Retrieved from http://dhsprogram.com/pubs/pdf/MR12/MR12.pdf

Schoumaker, B. (2014b). The crisscross method to evaluate data quality in fertility surveys. Presented at the PAA Meeting, Boston.

Zlotnik, H., & Hill, K. (1981). The Use of Hypothetical Cohorts in Estimating Demographic Parameters under Conditions of Changing Fertility and Mortality. *Demography*, *18*(1), 103–122. http://doi.org/10.2307/2061052