

# Inconsistencies in reports on siblings and children in survey data used to estimate mortality

Bruno Masquelier\*

Paper to be presented at the 2013 IUSSP Conference

## Abstract

### *Background*

In countries where the registration of deaths remains incomplete, the main sources of data on mortality are retrospective reports collected on the survival of close relatives. Birth and sibling histories included in household surveys are known to be plagued by recall errors but few methods exist to detect these errors.

### *Objectives*

This paper introduces two simple approaches to assess the consistency between sibling histories and reports on the fertility of the previous generation. The first is a comparison between the average size of sibships and the mean number of children ever born (CEB) to women of the previous generation, at the aggregate level. The second approach consists in linking at the individual level sibling histories reported by young women aged 15-18 with birth histories of their mothers when they both live in the same household.

### *Results*

When applied to Demographic and Health Surveys (DHS), the first approach indicates that a large proportion of siblings is omitted; on average the reported sibsizes are about 15 percent lower than expected from data on CEB. Discrepancies between reported and expected sibsizes are larger in sub-Saharan Africa than in other regions and they increase with the age of respondents. The proportions of brothers who died reported by sisters are lower than the corresponding proportions of deceased sons reported by mothers, except when restricting the analysis to the adult ages. The second approach highlights frequent inconsistencies between mother's and daughter's reports on members of the same family. Daughters tend to report fewer siblings than expected from their mother's birth history, but more adult deaths.

### *Conclusions*

Sibling histories are inadequate to estimate child survival, as are birth histories to investigate adult mortality. Even for the adult ages, survey reports on sibling survival are likely to under-estimate risks of dying. Linking sisters interviewed in the same household could help performing additional consistency checks and developing adjustments.

---

\*Catholic University of Louvain (UCL), Belgium - Associate researcher at INED, France.

## 1 INTRODUCTION

Survey data on the survival of siblings are a key source of mortality estimates in many countries where civil registration systems remain ineffective or incomplete. These data are now widely used to reconstruct trends in maternal mortality (Wilmoth et al. 2012, Hogan et al. 2010), as well as trends in all-cause adult mortality (Reniers et al. 2011). For example, in the Global Burden of Disease 2010 Study, sibling histories were the main source of information on adult survival in sub-Saharan Africa (Wang et al. 2012). Several survey programs have collected sibling histories, such as the World Health Survey by the World Health Organization (WHO), Reproductive Health Surveys of the U.S. Centers for Disease Control and Prevention (CDC), and Demographic and Health Surveys (DHS). DHS are nationally representative surveys focusing on fertility, family planning, HIV/AIDS and child and maternal health. Since 1989, a “maternal mortality” module has been included in more than 100 DHS surveys. This module is an equivalent of “full birth histories” (FBH) for the survival of siblings. In FBH, women of reproductive age provide an exhaustive list of all their children, and report their sex, date of birth, survival status, current age for surviving children and age at death if deceased. Additional questions are asked to locate surviving children in the roster of household members and to identify multiple births. Likewise, in “full sibling histories” (FSH), women aged 15-49 are asked to list all her siblings born to the same mother. Questions are also asked on their gender, survival status, current age when surviving, or ages at death and years since death when deceased. Some additional questions are aimed at identifying pregnancy-related deaths. Therefore, both FBH and FSH provide deaths distributed by age and calendar year with the corresponding person-years of exposure, allowing for the calculation of annual age-specific mortality rates (Masquelier et al. forthcoming).

Several studies have focused on the quality of FBH. They highlight in particular a problem of omissions of recent births or transference of some births out of a time window defined by the five years before the survey, in order to avoid a series of questions about the most recent births (Schoumaker 2009, Johnson et al. 2005). Less attention has been devoted to the quality of sibling histories.

Some studies evaluated mortality rates based on sibling survival data in comparison with estimates of the World Population Prospects (WPP) or with WHO life tables (Gakidou et al. 2004, Reniers et al. 2011, Stanton et al. 2000). They all concluded that sibling histories tend to underestimate adult mortality. However, it is difficult to draw definitive conclusions from such comparisons in absence of a gold standard, as mortality rates used as a reference can themselves be biased. This is especially the case in countries affected by HIV/AIDS, for which a complex modelling of the epidemic is needed. This can explain why the consistency between DHS and UN estimates of adult mortality varies from one region to another. For example, sibling histories provide very low levels of mortality in the Sahel, in comparison to those of the UN, while the two sets agree more in Eastern and Southern Africa (Reniers et al. 2011).

A second approach to assessing data quality was taken by Helleringer et al. (forthcoming). These authors linked retrospective siblings’ survival histories with data from

a demographic surveillance site in southeastern Senegal (Bandafassi). They showed that omissions of siblings are frequent, especially of deceased siblings (9%) and siblings who migrated out of the study area (17%). Age were also misreported, with a tendency to under-estimate the sisters' ages, particularly above age 45. This study contributes to a better understanding of the different types of recall errors in FSH, but it remains limited in scope as it is based only on one particular rural setting.

Thirdly, Stanton et al. (2000) performed internal consistency checks of FSH, focusing on the completeness of reports, the sex ratios of siblings, and the plausibility of the sizes of sibships (also referred to as sibsizes). Using 14 DHS conducted between 1989 and 1995, they showed that the data on reported events was complete in most surveys, and presented no strong indications of differential quality by sex of siblings or time period prior to the survey. However, Stanton et al. (2000) identified a few problems, including an abnormally high percentage of deaths located at 5, 10 or 15 years before the survey, and a concentration of deaths in the years immediately preceding the survey. This suggests that deaths that occurred in the distant past are disproportionately omitted. This was subsequently confirmed by Timaeus and Jasseh (2004) and Obermeyer et al. (2010). Since some countries have collected several sets of FSH, it is now possible to compare mortality rates estimated for the same country and the same period based on data from different surveys. Using this strategy, Masquelier et al. (forthcoming) showed that in sub-Saharan Africa, the completeness of death reporting declines below 75% when the reference periods extent in the past by more than six years from the survey (except for 10-11 years before the survey, because of heaping on 10). It is possible to adjust mortality rates accordingly, but this concerns only the relative under-reporting of deaths - compared to recent periods. The absolute level of omissions could be even higher. Moreover, estimating the relative under-reporting requires pooling several surveys together and it is therefore not possible to evaluate the quality of each survey taken individually. It is however important to assess whether the extent of omissions varies from one country to another, and potentially across surveys as well. It would also be useful to identify the categories of respondents that provide the least reliable information to improve the data collection or to develop adjustment procedures.

Building on the work of Stanton et al. (2000), this paper introduces two simple approaches to evaluate the quality of sibling data, both based on an assessment of the consistency between sibling histories reported by one generation of women and data on fertility of the previous generation.

1. First, the average sibsize reported in DHS is compared - at the aggregate level - with the mean number of children ever born (CEB) reported in previous censuses and surveys by women who have completed their reproductive life. This allows computing an "expected" number of siblings, an estimate that cannot be provided by any other source. This approach is adopted here to test two hypotheses. One is that the magnitude of omissions varies by region, with more frequent omissions in Western Africa. In this region, sibships are larger and also more complex, due to high levels of fertility and a higher incidence of polygamy. The second hypothesis is that older respondents disproportionately omit siblings. This has already been suggested by

Stanton et al. (2000). They observed in several surveys that the sibsize was invariant or decreasing with respondents' increasing age. To the authors, this pattern seemed inconsistent with fertility declines that occurred in most developing countries over the last decades.

Our approach based on CEB data will confirm that older respondents tend to omit slightly more siblings than younger respondents. It will also show that omissions are not confined to reports by older respondents; on average, the sibsizes reported in DHS are about 15 percent lower than expected from CEB data. These omissions seem more common in sub-Saharan Africa than in other developing regions, and particularly pronounced in surveys conducted in Western Africa. However, they seem to concern especially brothers and sisters who died in childhood, and therefore do not necessarily bias estimates of adult mortality.

2. The second approach differs from the first in that it is conducted at the individual level within the same households. Between ages 15 and 18, some women interviewed in DHS provide information on their sibships and also the line number of their mother, if the latter live in the same household. When this mother is under age 50, she provides a *birth history* which can be compared with the *sibling history* of her daughter.

This second approach will confirm that daughters report fewer siblings than expected from their mother's birth history, but more adult deaths.

After presenting these two approaches, this paper will discuss the implications of our results for the estimation of adult mortality. Refinements of the survey design will also be suggested.

## 2 COMPARISONS BETWEEN SIBSIZES AND LIFETIME FERTILITY OF WOMEN OF THE PRECEDING GENERATION

### 2.1 Data and method

The first part of this analysis is based on all available DHS including a module on maternal mortality. This corresponds to 104 surveys covering 50 developing countries. Although FBH have sometimes been collected from men's reports (Merdad et al. 2013), we focus here on data provided by women aged 15-49.

The mean sibsize is computed for each five-year age group (from age 15 to 49), including brothers and sisters as well as the respondents. Stanton et al. (2000) mentioned the DHS surveys carried out in Madagascar in 1992 and Senegal in 1992-93 as egregious examples of the decline of sibsizes with the respondent's age. This is illustrated in Figure 1, and supplemented with others surveys from these two countries. In 1992 and 1992/3, the sibsizes decrease rapidly with the age of respondents; women aged 45-49 report about 20% less siblings than women aged 20-24. This decline with age does not concern all surveys; in Madagascar in 2009, the opposite pattern is observed.

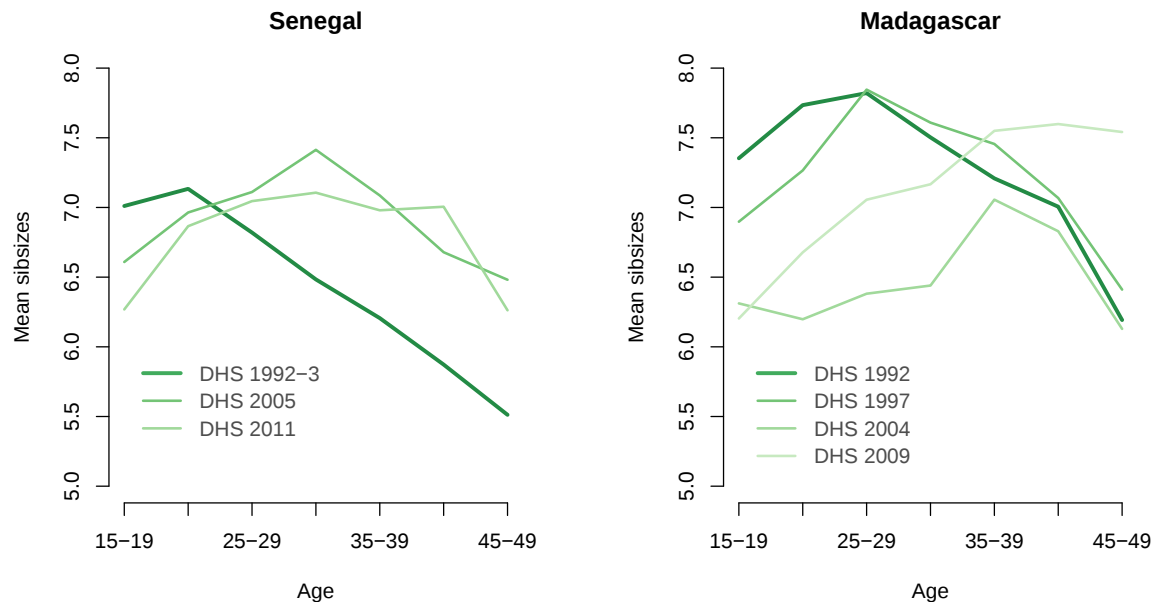


Figure 1. Average parity of respondents' mother by respondents' age - DHS conducted in Senegal and Madagascar

Comparing several surveys for a given country is seldom done but it can already highlight some inconsistencies. It is obvious that in Madagascar, omissions of siblings were more pervasive in 2004 than in other surveys. But it is also needed to introduce data on past trends in fertility. Indeed, older respondents aged 45-49 in 1993 were born around 1945, and their lower sibsize could reflect an rise in fertility between 1940 and 1970, rather than a tendency to omit more siblings. Likewise, particular care is needed when comparisons imply the mean sibsizes of respondents aged 15-19, as their mother may still be their reproductive period.

Preston (1976)<sup>1</sup> showed that the average size of sibships ( $F$ ) (including *ego*) is a function of the mean number of children ever born to the previous generation ( $G$ ) and the variance of this number ( $\sigma^2$ ), such that  $F = G + \sigma^2/G$ . If all women had the same number of children, the variance would be zero and all sibsizes would be equal to the number of live births. Once a variation is introduced in fertility rates, the average sibsize is inevitably higher than the mean number of CEB. This is simply because larger sibships are more frequently mentioned when children are asked about their sibship. An example is provided in Table 1, in which 100 women who completed their reproductive lives are distributed by parity according to a Poisson distribution (with parameter  $\lambda = 4$ ). If mortality does not vary with family size, the probability of selecting at random a mother who gave birth to three children is 0.24, while the probability of selecting at random a child who lived in a family of three children is  $3 \times 0.24$ . In this configuration, the mean number of CEB is 3.35, the variance of this completed fertility is 2.81, while the mean sibsize is 4.19, which

<sup>1</sup>? made the same demonstration at about the same period.

corresponds to  $3.35 + 2.81/3.35$ .

| 0                       | 1                                                                                         | 2  | 3  | 4  | 5  | 6 | 7 | 8 |
|-------------------------|-------------------------------------------------------------------------------------------|----|----|----|----|---|---|---|
| 5                       | 6                                                                                         | 21 | 24 | 21 | 11 | 9 | 2 | 1 |
| Mean CEB ( $G$ ):       | $0 \times 0.05 + 1 \times 0.06 + 0.21 \times 2 + \dots = 3.35$                            |    |    |    |    |   |   |   |
| Var CEB ( $\sigma^2$ ): | $(0.05 \times 0^2 + 0.06 \times 1^2 + 0.21 \times 2^2 + \dots) - G^2 = 2.81$              |    |    |    |    |   |   |   |
| Mean sibsize ( $F$ ):   | $(0.05 \times 0^2 + 0.06 \times 1^2 + 0.21 \times 2^2 + \dots)/G = G + \sigma^2/G = 4.19$ |    |    |    |    |   |   |   |

Table 1. Example of calculating the average number of brothers and sisters from the number of children ever born

It is therefore possible, once the cohort of mothers who gave birth to the women interviewed in DHS is identified, to recalculate the “expected” sibsize of the respondents, based on the mean and variance of the number of CEB reported by their mothers. To identify this cohort, it is necessary to have information on the age difference between mothers and offspring, that is to say the mean age at childbearing (MACB)<sup>2</sup>. This age varies little between countries or over time (between 25 and 30 years) and it tends to decrease during the fertility transition (Bongaarts 1999). Trends in MACB depends mainly on two factors: the delaying of births by young mothers (e.g. due to a rise in ages at first marriage) and the limitation of births by older mothers. This indicator can be estimated by taking the weighted average of ages of women tabulated by the number of births they had in the last 12 months. For the countries considered here, estimates of MACB were collected from DHS and UNICEF’s MICS surveys, the Demographic Yearbook of the United Nations (UN 1997; 2009), some other national demographic surveys and the samples of census micro-data distributed in the IPUMS database<sup>3</sup>. They are displayed as dots in Figure 2 (left plot). Some countries, such as Peru (which is singled-out), have many estimates, but trends are much less documented in others, such as Eritrea and Angola. To obtain a value for each country ( $j$ ) and each year ( $t$ ), a linear mixed model is used here<sup>4</sup>. In this model, the constant and the rate of decline in MACB are allowed to vary by country:

$$\begin{aligned}
 \text{MACB}_{ij} &= \beta_0 + u_{0j} + \beta_1 t_{ij} + u_{1j} t_{ij} + e_{ij} \\
 u_{0j} &\sim N(0, \sigma_{u0}^2) \\
 u_{1j} &\sim N(0, \sigma_{u1}^2) \\
 e_{ij} &\sim N(0, \sigma_e^2)
 \end{aligned} \tag{1}$$

The advantage of a mixed model is that the resulting predicted values are close to the average in countries where there are few estimates or when they fluctuate greatly; they “stick” more to the estimates of a specific country where the latter are more numerous and

<sup>2</sup>This indicator is often confused with the *mean age of the fertility schedule*, which is obtained by taking the average of women’s ages weighted by fertility rates. While the MACB is influenced by the age structure, the mean age of the fertility schedule is not; the gap between both indicator can therefore be quite large.

<sup>3</sup>Minnesota Population Center. Integrated Public Use Microdata Series, International: Version 6.1 [Machine-readable database]. Minneapolis: University of Minnesota, 2013.

<sup>4</sup>The package nlme from the R statistical software is used here (Pinheiro et al. 2013).

more regular (Gelman and Hill 2007).

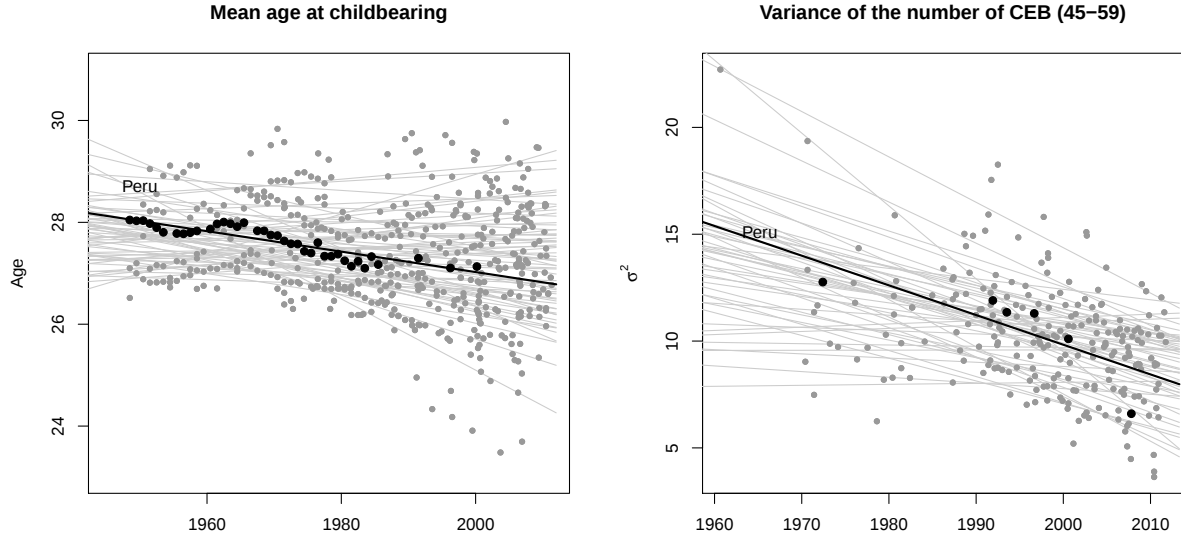


Figure 2. Trends in the mean age at childbearing and the variance of completed fertility of women aged 45 to 49 (for 50 developing countries covered by DHS sibling histories).

The same model is used to reconstruct trends in the variance of completed fertility. This variance is calculated from women aged 45-49 because this is the age group for which there is the greatest amount of data, since women older than 50 years are not interviewed in DHS. In addition to estimates from DHS (159 surveys), IPUMS samples (44 censuses) as well as census reports and other national surveys (60 operations) are used here. As the mean age at childbearing, the variance of completed fertility tends to decrease with decreasing fertility (Figure 2 - right plot).

Data on the number of children ever born come from the same sources but are more abundant (some reports of surveys or censuses publish the average parities without giving the full distribution). A total of 369 operations are available for the 50 countries considered here. Only the estimates of completed fertility of women ages 45-49, 50-54, 55-59, and 60-64 are retained.

## 2.2 Results

CEB data are usually considered of poor quality when reported by post-reproductive age women, with a tendency to under-report children (Nations Unies 1983)<sup>5</sup>. Feeney (1991)

<sup>5</sup>Another problem arises in some censuses or surveys when childless women are confused with women whose number of CEB is not reported. This is the case for example if the enumerators indicate a dash (-) in the appropriate box on the questionnaire. El-Badry (1961) developed a method to adjust the number of CEB accordingly. I applied this method in cases where the distribution of women by age and CEB was available (eg with IPUMS data), when the proportion of missing responses was larger than 2% and when

showed, however, that in many countries, trends in CEB were remarkably consistent from one census or survey to another. He suggests ‘time-plotting’ CEB data, taking as value for the x-axis the year in which women have reached the mean age at childbearing. Figure 3 (left plot) provides an example with the case of Peru, where a census was held in October 2007. Women aged 45-49 enumerated at that time were born in April 1960<sup>6</sup>. In August 1987, they reached 27.3 years, the mean age at childbearing estimated for the country at this period. This is retained as the reference period to present their mean number of CEB.

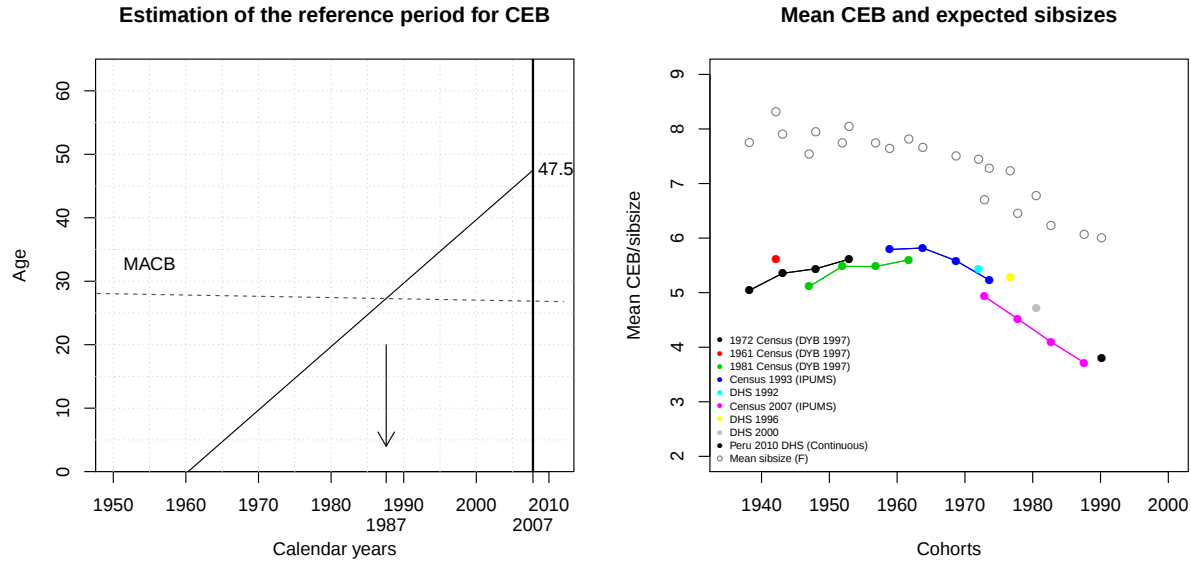


Figure 3. Example of calculation of the reference period (right graph), TCFR and sizes expected in Peru (left graph) siblings

By doing so for the other three age groups until 60-64 years and for each survey or census carried out in Peru for which CEB data are available, one obtains the trend shown in the right plot of Figure 3. The good consistency between mean CEB reported throughout the period is noteworthy. It is possible that all estimates are plagued by under-reporting in the same proportion, but the overall consistency at least shows that the effect did not worsen with increasing age. The mean number of CEB corresponding to births that occurred on average between 1940 and 1965 ranged between 5 and 6, then they dropped to 3.5 in the most recent periods. Estimates derived from DHS are slightly higher than those based on censuses, probably because of the additional questions on birth histories in DHS. Combined with the variance of the number of CEB presented above, these estimates can be used to estimate the “expected” sizes of sibships (displayed with white dots). The sibsize of cohorts born until 1965 was close to 8, and then it decreased to reach 6 for the cohorts born around 1990. A comparable number of brothers and sisters (including the respondents) should be reported in DHS, or even a larger number if siblings are better reported than are children.

the conditions necessary for this adjustment were met.

<sup>6</sup>They are assumed to be evenly distributed between 45 and 50 years.

Unsurprisingly, this is seldom the case. Figure 4 presents, for six countries, the expected sibsizes (in dark gray), along with sibsizes reported in DHS (solid lines). Various patterns are observed. In Peru, the reported sibsizes are larger than the number of CEB, which is normal, but they remain below the “expected” sibsizes. The difference amounts to 1.34 sibling in the 1992 survey, against 0.59 in the following two surveys. In the Dominican Republic, reported sibsizes are very close to expected sibsizes, and successive surveys are relatively consistent. This configuration reflects a high quality of data but it is observed in less than a dozen countries, located mostly in Asia and Latin America (such as Brazil, the Philippines, Nepal and Jordan). The pattern observed in Zimbabwe, Tanzania and Malawi is the most common in sub-Saharan Africa: the average sibsize is lower, by a little more than one sibling, than the expected sibsize. Finally, in some surveys, a very implausible situation is observed, with mean sibsizes lower than mean number of CEB, as in Malawi in 2004 and Nigeria in 1999, a survey known for its poor quality (Pullum 2008).

Comparing the reported sibsizes with data on lifetime fertility allows to qualify the statement by Stanton et al. (2000); in some cases, a rise in sibsizes with respondent’s age might not reflect a problem of data quality but stems from past trends in fertility. However, our analysis does confirm that omissions are more frequent among older respondents. If one assumes that the CEB data are accurate, the proportion of brothers and sisters who have not been reported reaches 16% on average over all surveys<sup>7</sup>. This proportion varies from 14% among respondents ages 25 to 34 to 20% among women aged 45-49. It is slightly higher among respondents aged 15-19 (17%), compared to those aged 20-24, probably because their sibships are not yet complete.

Regional variations are more pronounced. For respondents above age 20, the discrepancies between ‘expected’ and reported mean sibsizes reaches 10% on average in DHS in Latin America and the Caribbean, against 15% in South Africa and 22% in West Africa.

---

<sup>7</sup>In order to quantify the differences between reported and expected sibsizes, it is necessary to interpolate between expected sibsizes. A cubic spline function is used here (Faraway 2006).

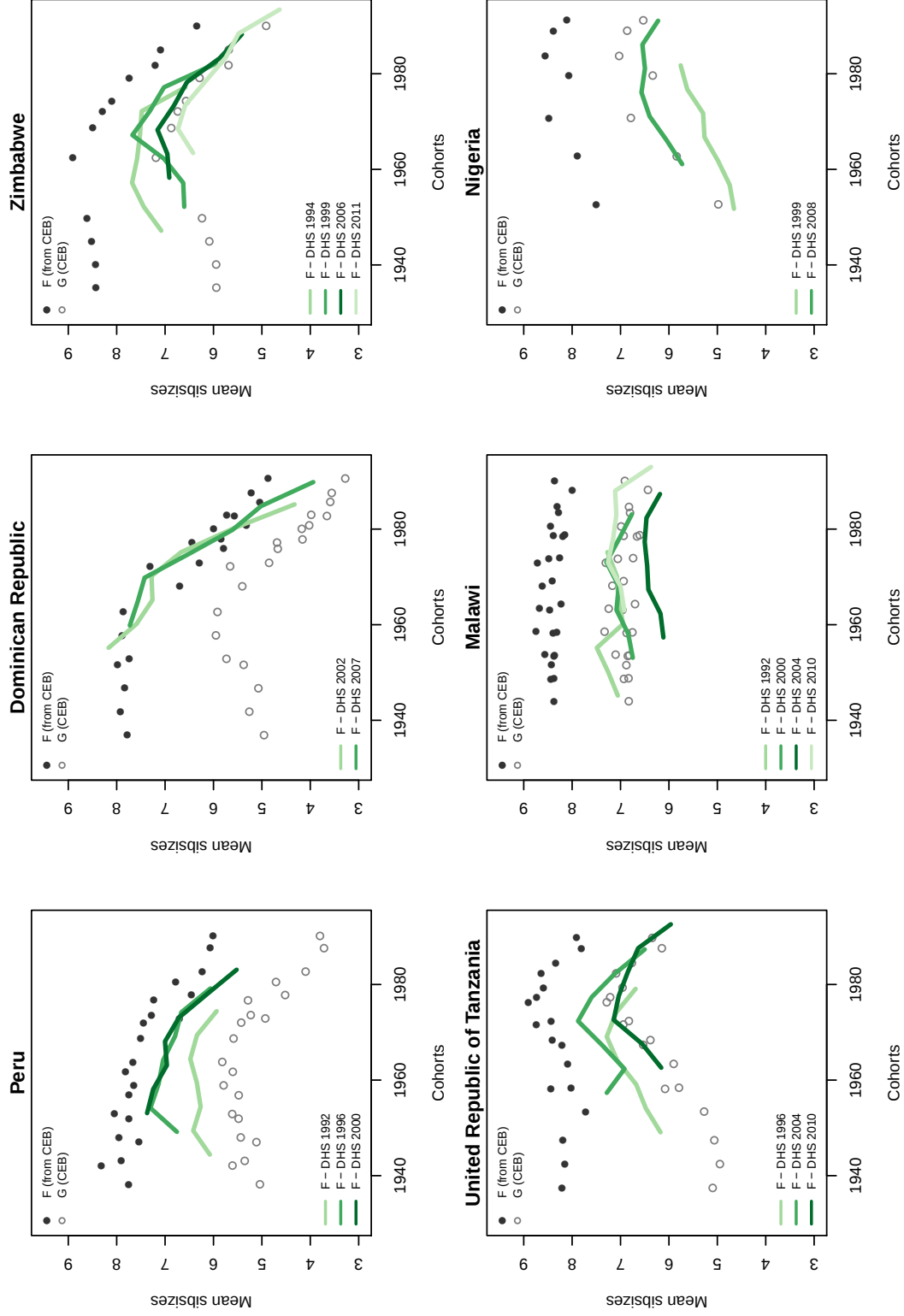


Figure 4. Comparison of the expected sizes with data reported in the DHS sizes siblings Peru, Zimbabwe, Dominican Republic, Tanzania, Malawi, Nigeria.

Omissions of siblings do not necessarily result in an under-estimation of mortality, unless the deceased are disproportionately omitted. Again, comparing sibling histories with reports by women of the preceding generation is instructive. In a given survey, respondents aged 15-19 provide information on cohorts of children born to mothers who are about 45 years old, as the mean age at childbearing is generally around 27 years. The proportion of deceased brothers can therefore be compared with the proportion of deceased sons in the same DHS. Here, only the information related to brothers is retained in order to avoid biases related to the fact that respondents aged 15-19, who are all women, are also all surviving.

The left plot in Figure 5 displays, on the x-axis, the proportion of deceased brothers, as reported by their sisters aged 15 to 19, compared with, on the y-axis, the proportion of deceased sons of women aged 17.5 + MACB (based on interpolations within the surrounding age groups). In almost all surveys, proportions of deceased sons are higher than proportions of deceased brothers, even though both types of proportions emanate from the same surveys and refer to the same cohorts (but not necessarily to the same individuals as the comparison is made at the aggregate level). Of course, these differences are not solely attributable to differential omissions; they are probably partly influenced by selection biases. For example, a sibships in which all siblings died can still be reported by mothers but not by sisters. By contrast, the fact that some women aged 45 may not have reached their completed fertility does not distort this comparison as both proportions refer to children born at the time of the survey.

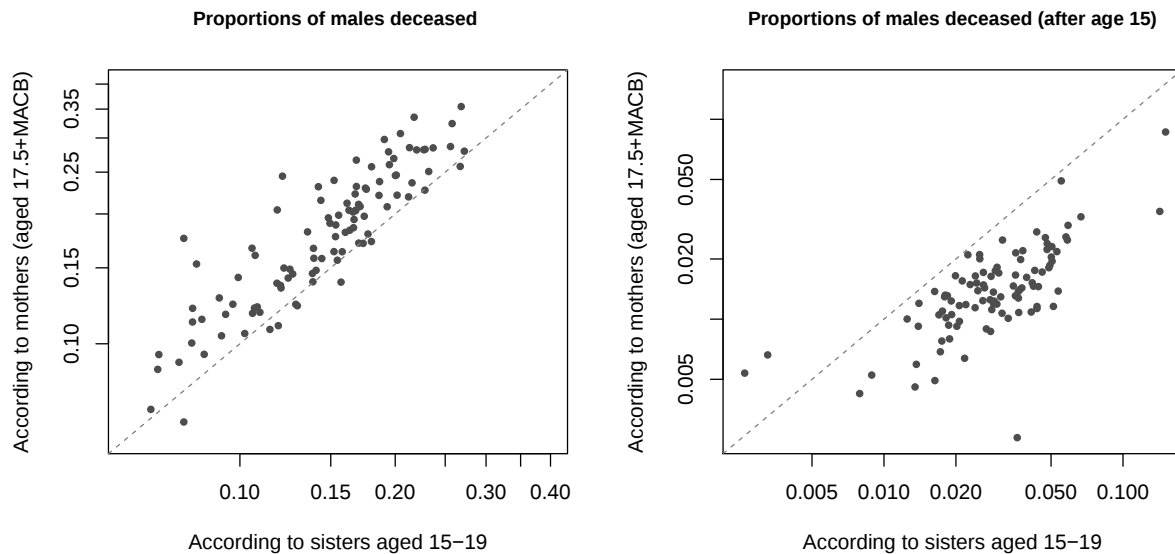


Figure 5. Proportion of males deceased, according to their sisters (aged 15-19) and their mothers (aged 17.5 + MACB).

When only brothers or sons who survived to age 15 are retained, the opposite pattern is observed; proportions of brothers who died in adulthood are higher than proportions of sons who died above age 15 (Figure 5, right plot).

### 3 LINKING BIRTH HISTORIES AND SIBLING HISTORIES AT THE HOUSEHOLD LEVEL

#### 3.1 Data and methods

Since the early 1990s, DHS have included questions on the survival status of biological parents and their residence status for children enumerated in the roster of household members. These questions aim at identifying orphans, and analysing fostering practices and co-residence with surviving parents (Monasch and Boerma 2004). When the mothers reside in the same household as the children, the interviewer writes down their line number. In several DHS surveys, questions on parental survival and residence have been extended to children under age 18 (Palermo and Peterman 2009). Since young women aged 15 to 18 are eligible to respond to the maternal mortality module, they provide data on both their siblings and parents. If they reside with a mother who is also eligible to the individual questionnaire (i.e. if she is under age 50), the sibling history they report can be compared with the birth history provided by their mother.

This approach is used here to confirm that sisters tend to report smaller family sizes than do mothers, but a larger number of adult deaths. A comparable analysis (restricted to women who are household head) has recently been performed by Merdad (2013) in her dissertation, in an attempt to test whether sibling histories can be a valuable source of data on child mortality.

#### 3.2 Results

Table 2 displays, for 22 DHS conducted in Sub-Saharan Africa, the number of mother-daughter pairs retained in this analysis (column 3), the percentage of sibships for which the number of *children* ever born reported by mothers is higher (4) or lower (5) than the number of *siblings* reported by their daughters, and the percentage of inconsistencies on the number of survivors (6) and deaths (7). There is a considerable variation in the extent of concordance in the reported sibship sizes. The percentages of inconsistent reports range from 4% in the Rwanda 2005 survey to 25% in Nigeria in 1999. Mother-daughter inconsistencies are predominantly attributable to lower sibship sizes reported by daughters. On average, daughters reported 0.13 fewer siblings than mothers and this difference reached 0.57 siblings in the 1999 Nigeria survey. There is therefore a higher consistency between birth and sibling histories in terms of family sizes than was observed in the previous section; this could be due to the fact that the same information is collected twice by the interviewers. Once the link between mothers and daughters is established, the interviewers could potentially copy the information entered in the mother's birth history. It is also possible that some relatives have helped in providing information for both modules.

As expected, the discrepancies between daughters' and mother's reports are generally larger for deceased siblings. However, much of the agreement between mothers and daughters on the number of deaths is attributable to their relative rarity. A more precise view is obtained by restricting the analysis to mother-daughter pairs with at least one mention of a death. Inconsistencies are observed in around 25% of the 4638 families with at

| Survey       | Year | n    | CEB > sib. | CEB < sib. | CEB $\neq$ sib. | CEB $\neq$ sib. |
|--------------|------|------|------------|------------|-----------------|-----------------|
|              |      | (3)  | (4)        | (5)        | Surv. (6)       | Deaths (7)      |
| Rwanda       | 2005 | 731  | 0.02       | 0.02       | 0.04            | 0.05            |
| Sao Tome     | 2009 | 146  | 0.03       | 0.03       | 0.04            | 0.05            |
| Chad         | 2004 | 436  | 0.03       | 0.01       | 0.03            | 0.04            |
| Congo        | 2005 | 332  | 0.05       | 0.07       | 0.05            | 0.07            |
| Madagascar   | 2009 | 1034 | 0.05       | 0.01       | 0.04            | 0.04            |
| Tanzania     | 2004 | 583  | 0.05       | 0.02       | 0.03            | 0.06            |
| Malawi       | 2004 | 496  | 0.06       | 0.01       | 0.03            | 0.07            |
| Benin        | 2006 | 762  | 0.07       | 0.02       | 0.06            | 0.08            |
| Zambia       | 2007 | 383  | 0.07       | 0.02       | 0.04            | 0.09            |
| DR Congo     | 2007 | 485  | 0.08       | 0.04       | 0.07            | 0.10            |
| Mozambique   | 2003 | 513  | 0.08       | 0.00       | 0.02            | 0.08            |
| Lesotho      | 2004 | 334  | 0.09       | 0.03       | 0.06            | 0.11            |
| Lesotho      | 2009 | 342  | 0.10       | 0.07       | 0.09            | 0.12            |
| Nigeria      | 2008 | 1459 | 0.10       | 0.02       | 0.07            | 0.11            |
| Guinea       | 1999 | 92   | 0.11       | 0.02       | 0.08            | 0.12            |
| Liberia      | 2007 | 311  | 0.11       | 0.04       | 0.10            | 0.10            |
| Namibia      | 2007 | 376  | 0.11       | 0.06       | 0.09            | 0.10            |
| Uganda       | 2006 | 513  | 0.12       | 0.05       | 0.12            | 0.09            |
| Zimbabwe     | 2006 | 396  | 0.13       | 0.02       | 0.09            | 0.11            |
| Swaziland    | 2006 | 281  | 0.19       | 0.03       | 0.08            | 0.16            |
| Sierra Leone | 2008 | 181  | 0.20       | 0.10       | 0.24            | 0.15            |
| Nigeria      | 1999 | 176  | 0.22       | 0.03       | 0.22            | 0.13            |

*Table 2. Percentage of mother-daughter pairs with discrepant reports on the number of children/siblings ever born, broken down by survival status*

least one child deaths. As shown in the previous section, sisters typically report less child deaths than do mothers. Differences in reporting of recent adult deaths (that occurred in the 10 years prior to the survey) are much more frequent; they concern about half of the mother-daughter pairs in which at least one recent adult death was mentioned. In this case, mothers report less adult deaths than do sisters.

## 4 CONCLUSION

This paper introduced two simple approaches to evaluate the consistency between sibling histories and data on children ever born. The first approach, at the aggregate level, reveals that a significant share of siblings is omitted in the DHS; the reported sibsizes are on average 15% lower than expected based on CEB data. The hypothesis that older women provide data of a lesser quality is verified, but omissions plague reports from all age groups. These omissions are more common in sub-Saharan Africa than in other developing regions.

| Survey       | Year | Child deaths |          |          | Recent adult deaths |          |          |
|--------------|------|--------------|----------|----------|---------------------|----------|----------|
|              |      | n            | sib<moth | sib>moth | n                   | sib<moth | sib>moth |
| Rwanda       | 2005 | 411          | 78       | 19       | 97                  | 4        | 66       |
| Sao Tome     | 2009 | 46           | 11       | 4        | 10                  | 1        | 6        |
| Chad         | 2004 | 259          | 26       | 8        | 30                  | 3        | 13       |
| Congo        | 2005 | 134          | 18       | 15       | 10                  | 1        | 5        |
| Madagascar   | 2009 | 349          | 45       | 9        | 27                  | 4        | 3        |
| Tanzania     | 2004 | 314          | 33       | 9        | 17                  | 1        | 4        |
| Malawi       | 2004 | 278          | 40       | 8        | 21                  | 1        | 5        |
| Benin        | 2006 | 368          | 57       | 15       | 19                  | 2        | 6        |
| Zambia       | 2007 | 187          | 39       | 6        | 20                  | 1        | 7        |
| DR Congo     | 2007 | 239          | 62       | 19       | 17                  | 4        | 4        |
| Mozambique   | 2003 | 256          | 45       | 5        | 21                  | 3        | 2        |
| Lesotho      | 2004 | 104          | 28       | 10       | 21                  | 4        | 5        |
| Lesotho      | 2009 | 80           | 32       | 10       | 7                   | 1        | 1        |
| Nigeria      | 2008 | 703          | 168      | 30       | 49                  | 5        | 18       |
| Guinea       | 1999 | 60           | 11       | 2        | 2                   | 0        | 1        |
| Liberia      | 2007 | 139          | 33       | 8        | 15                  | 2        | 6        |
| Namibia      | 2007 | 92           | 27       | 7        | 13                  | 1        | 2        |
| Uganda       | 2006 | 302          | 36       | 32       | 21                  | 4        | 4        |
| Zimbabwe     | 2006 | 89           | 41       | 7        | 19                  | 4        | 4        |
| Swaziland    | 2006 | 88           | 45       | 2        | 20                  | 5        | 5        |
| Sierra Leone | 2008 | 78           | 23       | 14       | 11                  | 4        | 6        |
| Nigeria      | 1999 | 62           | 24       | 2        | 1                   | 0        | 1        |

Table 3. Total mother-daughter pairs with at least one report of child death or recent adult death, and distribution of discrepant reports on these numbers

In particular, several surveys carried out in West Africa are characterized by very large discrepancies between reported and expected sibsizes: reported sibsizes are 37% lower than expected in Sierra Leone in 2008, 35% in Nigeria in 1999, 31% in Liberia in 2007 and 25% in Niger 1992. This does not mean that the under-estimation of adult mortality is larger in this region, because omissions seem to concern especially brothers and sisters who died in childhood - and child mortality is the highest in West Africa.

What are the implications of these results for the estimation of *adult* mortality? There are currently three approaches for estimating adult survival from sibling data. Hill and Trussell (1977) first developed an indirect technique, which does not discriminate between deaths that occurred in childhood and those that occurred in adulthood. Results presented here suggest that it will lead to severe under-estimation of mortality. The indirect method developed by Timaeus et al. (2001) is to be preferred. The latter applies to surveys or

censuses in which some additional questions are asked about the number of siblings who survived to age 15. However, it will mix deaths that occurred in recent and distant periods, which makes it sensitive to the under-reporting of adult deaths very distant from the survey. By contrast, the direct approach is immune to omissions of siblings who died in childhood or many years prior to the survey (Rutenberg and Sullivan 1991). It requires that questions on ages at surveys and ages at deaths are systematically included in the questionnaires. In the absence of complete vital registration systems, full birth histories should be collected in all national surveys focusing on health and mortality. This is currently the case of DHS, but not MICS surveys, where only a summary sibling history is asked.

Improvements could also be made to the standard “maternal mortality module” used in DHS, by identifying the sisters living in the same household (thus providing information on the same family several times). This would allow various data quality checks and would help develop corrections for under-reporting. Identifying sisters that are eligible to the individual questionnaire has been tried in some DHS with a question asked at the end of the maternal mortality module (Niger 1992, Senegal 1992-3), but the data are incomplete or the field is left blank. Alternatively, eligible women who have the same mothers could be identified when completing the household questionnaires.

In the published reports of DHS, the average parity of respondents’ mother by respondents’ age is routinely used as an indicator of data quality. The lack of monotonic increase of sibsize with age is seen as a sign of omissions. The analysis presented here indicates that this indicator is too crude. Insofar as they were constituted between 15 and 50 years before the survey, “expected” sibsizes do not necessarily decline with respondent’s age; invariance is expected in Malawi, for example (Figure 5). Taken alone, the sibsize is not informative. The comparative exercise proposed here is based on approximations, however. It implies for example that all women have given birth when they reached the mean age at childrearing. It also requires that there is no association between child mortality and fertility of mothers, or between the size of sibships and the adult survival. Moreover, the “expected” sibsize depends of course on the quality of reporting of CEB. As the CEB data are known to be themselves plagued by under-reporting, the extent of omissions of siblings could be even larger than estimated here.

## References

- J. Bongaarts. The Fertility Impact of Changes in the Timing of Childbearing in the Developing World. *Population Studies*, 53(3):277–289, 1999.
- M. A. El-Badry. Failure of Enumerators to Make Entries of Zero: Errors in Recording Childless Cases in Population Censuses. *Journal of the American Statistical Association*, 56(296):pp. 909–924, 1961.
- J. Faraway. *Extending the linear model with R*. Tylor & Francis, 2006.
- G. Feeney. Child survivorship estimation: Methods and data analysis. *Asian and Pacific Population Forum*, 5(2-5):51–55, 76–87, 1991.

- E. Gakidou, M. Hogan, and A. Lopez. Adult mortality: time for a reappraisal. *International Journal of Epidemiology*, 33, 4:710–717, 2004.
- A Gelman and J Hill. *Data analysis using regression and multilevel/Hierarchical models*. Cambridge University Press, 2007.
- S. Helleringer, G. Pison, M. Kanté, G. Duthé, and A. Andro. Reporting errors in survey data on adult mortality: results from a record linkage study in Senegal. *Demography*, forthcoming.
- K. Hill and J. Trussell. Further developments in indirect mortality estimation. *Population Studies*, 31(2):313–334, 1977.
- M. Hogan, K. Foreman, M. Naghavi, S. Ahn, M. Wang, S. Makela, A. Lopez, R. Lozano, and Ch. Murray. Maternal mortality for 181 countries, 1980-2008: a systematic analysis of progress towards Millennium Development Goal 5. *The Lancet*, 375(9726):1609 – 1623, 2010.
- K. Johnson, Sh. Rutstein, and P. Govindasamy. *The Stall in Mortality Decline in Ghana: Further Analysis of Demographic and Health Surveys Data*. Calverton, Maryland, USA: ORC Macro., 2005.
- B. Masquelier, G. Reniers, and G. Pison. Divergences in child and adult mortality trends in sub-Saharan Africa: evidence from survey data on the survival of children and siblings. *Population Studies*, forthcoming.
- L. Merdad. Investigating New Methods for Estimating Maternal and Child Mortality in the Developing World. Unpublished dissertation, Harvard School of Public Health, 2013.
- L. Merdad, K. Hill, and W. Graham. Improving the Measurement of Maternal Mortality: The Sisterhood Method Revisited. *PLoS ONE*, 8(4):e59834, 04 2013.
- R. Monasch and J. Boerma. Orphanhood and childcare patterns in sub-Saharan Africa: an analysis of national surveys from 40 countries. *AIDS*, 18(2):55–65, 2004.
- Nations Unies. *Manuel X: Techniques indirectes d’estimation démographique*. Population Division, Dept. of International Economic and Social Affairs, United Nations, New York, 1983.
- Z. Obermeyer, J. Rajaratnam, C. Park, E. Gakidou, M. Hogan, A. Lopez, and Ch. Murray. Measuring Adult Mortality Using Sibling Survival: A New Analytical Method and New Results for 44 Countries, 1974-2006. *PLoS Medicine*, 7(4):e1000260, 04 2010.
- T. Palermo and A. Peterman. Are Female Orphans at Risk for Early Marriage, Early Sexual Debut, and Teen Pregnancy? Evidence from Sub-Saharan Africa. *Studies in Family Planning*, 40(2):101–112, 2009.
- J. Pinheiro, D. Bates, S. DebRoy, D. Sarkar, and R Core Team. *nlme: Linear and Nonlinear Mixed Effects Models*, 2013. R package version 3.1-109.
- S. Preston. Family Sizes of Children and Family Sizes of Women. *Demography*, 13(1):105–114, 1976.

- Th. Pullum. An Assessment of the Quality of Data on Health and Nutrition in the DHS Surveys, 1993-2003. DHS Methodological Reports No. 6., Calverton, Maryland: Macro International Inc., 2008.
- G. Reniers, B. Masquelier, and P. Gerland. Adult Mortality in Africa. In R. Rogers and E. Crimmins, editors, *International Handbook of Adult Mortality*. Springer, 2011.
- N. Rutenberg and J. Sullivan. Direct and indirect estimates of maternal mortality from the sisterhood method. Proceedings of the Demographic and Health Surveys World Conference, August 5-7 1991, Washington, D.C., Vol. III. IRD/MACRO International Inc., Columbia , Maryland., 1991.
- B. Schoumaker. Stalls and Reversals in Fertility Transition in Sub-Saharan Africa. Real or Spurious ? Working paper SPED 30, Louvain-la-Neuve, 2009.
- C. Stanton, A. Nouredine, and K. Hill. DHS Maternal Mortality Indicators: an assessment of data quality and implications for Data Use. *Studies in Family planning*, 31,2:111–123, 2000.
- I. Timaeus and M. Jasseh. Adult mortality in Sub-Saharan Africa: evidence from demographic and health survey. *Demography*, 41, 4:757–772, 2004.
- I. Timaeus, B. Zaba, and M. Ali. Estimation of adult mortality from data on adult siblings. In B. Zaba and J. Blacker, editors, *Brass tacks: essays in medical demography*, pages 43–66. Athlone Press, Londres, 2001.
- UN. *Demographic Yearbook 1997 - Historical supplement*. Dept. of International Economic and Social Affairs, United Nations, 1997.
- UN. *Demographic Yearbook 2007*. Dept. of International Economic and Social Affairs, United Nations, 2009.
- Haidong Wang, Laura Dwyer-Lindgren, Katherine T Lofgren, Julie Knoll Rajaratnam, Jacob R Marcus, Alison Levin-Rector, Carly E Levitz, Alan D Lopez, and Christopher JL Murray. Age-specific and sex-specific mortality in 187 countries, 1970-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*, 380(9859):2071–2094, December 2012.
- J. Wilmoth, N. Mizoguchi, M. Oestergaard, L. Say, C. Mathers, S. Zureick-Brown, M. Inoue, and D. Chou. A New Method for Deriving Global Estimates of Maternal Mortality. *Statistics, Politics, and Policy*, Volume 3, Issue 2:2151–7509, 2012.

## Acknowledgements

I am grateful to Gilles Pison for constructive comments on this paper and to Patrick Gerland for providing CEB data for sub-Saharan Africa. This research was conducted thanks to a grant from the Belgian National Fund for Scientific Research (FNRS).