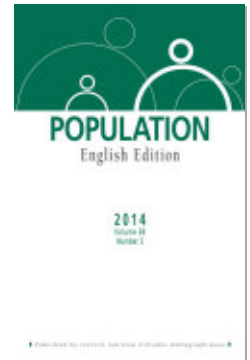# Estimating Age at First Union in Africa. Are Census and Survey Data Comparable?

Véronique Hertrich, Solène Lardoux, Karen George

Véronique HERTRICH* and Solène LARDOUX**

Estimating Age at First Union in Africa. Are Census and Survey Data Comparable?

Age at entry into union is a fundamental variable in the timing of conjugal and family life. Do censuses and population surveys yield the same estimates of age at first union? Should one of these sources be preferred over the other? For countries where it can be difficult to record ages and where the study of nuptiality is generally based solely on survey data, these questions deserve special attention. Véronique HERTRICH *and* Solène LARDOUX *compare estimates of age at marriage drawn from 450 data collection operations – both surveys and censuses – carried out in 55 African countries since the 1950s, and demonstrate that these two types of source each introduce biases that pull in opposite directions. These inconsistencies originate in reporting errors at the time of data collection and are particular to each type of source. So there is no reason to prefer surveys over censuses when analysing the timing of nuptiality; rather, the two sources should be used in tandem.*

The production of national population statistics, although late in getting off the ground (in many countries the first survey dates from the 1960s and the first census from the 1970s), has made considerable progress in Africa over recent decades. As part of this process, in addition to ten-yearly censuses, huge demographic research programmes have been established, from the World Fertility Surveys (WFS, 1975-1985) through to the latest rounds of Demographic and Health Surveys (DHS); there have also been more specific programmes, such as the League of Arab Nations' PAPCHILD (Pan Arab Project for Child Development) and PAPFAM (Pan Arab Project for Family Health) surveys and the MICS (Multiple Indicator Cluster Survey) studies under the guidance of UNICEF, as well as independent national surveys. As a result, a significant quantity of national demographic data is now available at the scale of the whole continent and for most African countries. In the 55 countries of Africa, a low

* Institut national d'études démographiques (INED), Paris.

** Université de Montréal, Département de démographie.

Correspondence: Véronique Hertrich, Institut national d'études démographiques, 133 boulevard Davout, 75980 Paris cedex 20, tel: +33 0 (1) 56 06 21 32, e-mail: hertrich@ined.fr

estimate[1] puts the number of censuses and national surveys carried out between 1950 and 2010 at over 500, an average of more than nine for each country.

Although some of the information collected by censuses and surveys is identical, these different data sources are rarely used in an integrated and systematic manner to examine long-term demographic trends at the scale of the African continent. This is certainly the case for nuptiality trends. Period estimates of age at first union can be obtained from both censuses and surveys by using the standard statistical table of marital status by sex and age. However, after the seminal work in this field in the 1980s (Lesthaeghe et al., 1989; van de Walle, 1993), most publications describing African nuptiality trends have used only a selection of the existing data.[2]

Why is the potential for comparative analysis offered by available statistical operations so neglected? Two main reasons can be suggested: first, data accessibility (i.e. use of data is restricted by difficulties in gaining access to it), and second, source comparability (the data collection protocols of the various sources are not uniform enough to guarantee that data from independent operations will be comparable).

For a long time, dissemination of census and survey results was often delayed and limited in scope, making analysis difficult. But the situation has changed considerably. Since the launch of the DHS, swift publication of results and easy access to data have become the norm for most African surveys. Census publications are available more quickly and easily on the websites of national statistical institutions, while the IPUMS project (Integrated Public Use Microdata Series) provides access to census microdata on a growing number of countries. Finally, with specific regard to nuptiality, the United Nations' database (2008, 2013) that brings together indicators drawn from several different sources for each country should now make it easier to develop systematic comparative approaches (Ortega, 2014).

But even with this increasing access to data, there is still the question of whether data from censuses and surveys are sufficiently comparable to be treated as a single corpus. Does the heterogeneity of data sources bias the reconstruction of long-term trends in age at first union? Are there consistent types of differences between the estimates produced and, if so, what causes them? Is it possible to determine whether censuses should be preferred over surveys, or vice-versa, to obtain high-quality estimates?

These questions will be addressed using two approaches. The first compares estimates of age at first marriage taken from censuses and surveys, drawing

---

(1) Based on the results of Gendreau and Gubry (2009) and Hertrich and Lardoux (2009).

(2) For instance, some publications describe trends by comparing period estimates from successive DHS surveys, while others focus on retrospective measures from one single survey, or compare estimates at two points in time (Garenne, 2004, 2014; Lloyd, 2005; Mensch et al., 2005, 2006; Shapiro and Gebreselassie, 2014; Tabutin and Schoumaker, 2004; Westoff, 2003).

on a pan-African database on nuptiality that brings together more than 450 censuses and surveys from the 55 countries of Africa. The second is based on analysis of a corpus of 15 MICS surveys, which recorded women's marital status on both a household and an individual questionnaire. After determining the frequency and the direction of differences between census-based and survey-based estimates, these secondary analyses of individual data will provide a more precise view of the mechanisms that give rise to such distortions. We will begin by examining the factors that may affect the quality and comparability of estimates of age at marriage drawn from survey and census data.

The terms "marriage" and "union" are used interchangeably here when discussing women's first unions, without reference to whether or not they have been formalized.

I. Should we expect censuses and surveys to yield different estimates of median age at first marriage?

Cross-sectional estimates: a means of avoiding retrospective reporting biases

In the absence of vital statistics, there are two main methods for measuring age at marriage within a population, using either respondents' retrospective reports (age at or date of marriage) or the proportions of never-married individuals by age recorded at a given point in time. Retrospective data are provided by most demographic surveys and can be used to estimate trends directly. However, their quality is limited not only by the recall errors common to all retrospective reporting but also by difficulties, particular to sub-Saharan Africa, in dating unions. For one thing, marriage processes involve various ceremonies and stages, and this leads to flexible and varying interpretations of the timing of entry into union (van de Walle, 1968; Mair, 1971; Meekers, 1992; Hertrich and Locoh, 1999; Antoine et al., 2009; Hertrich, 1998, 2007b; Lardoux, 2009). Furthermore, it remains difficult to determine a precise date or age in contexts where these concepts may only recently have come into use (Roger et al., 1981; Ewbank, 1981; Waltisperger, 1988), so there is a risk of recording normative, imprecise responses. Methodological studies have concluded that these retrospective data on age at marriage in Africa are of poor or, at best, middling quality (van de Walle, 1968, 1993; Lesthaeghe et al., 1989; Blanc and Rutenburg, 1990; Gage, 1995; Hertrich and Lardoux, 2009; Chae, 2011). Ron Lesthaeghe (1989) and Étienne van de Walle (1968, 1993) advocate rejecting them in favour of cross-sectional indicators.

Using cross-sectional data we can avoid the risks of mistaken interpretation and dating of past events by focusing on the structure of the population by sex, age and marital status at the time of the survey or census. Under the approach proposed by Hajnal (Hajnal, 1953; Tabutin and Vallin, 1975; United

Nations, 1984; Gubry, 1984), the series of proportions of never-married individuals by age can be equated with that of a theoretical cohort and summarized by a standard indicator such as mean age or median age at first marriage. Where entry into union is concentrated within a narrow age range, the indicator captures the nuptiality of the cohorts reaching those ages at the time of survey. In sub-Saharan Africa, where most women marry young, median age at first union is strongly correlated with the proportion of never-married women aged 15-19 and 20-24.[3] Another advantage of this method is that it can be applied to most data collection operations: marital status is generally recorded by censuses and surveys and published in a statistical table by sex and age group.

Key variables: age at time of survey and marital status

When median age at first marriage is calculated from period data, the constraints of retrospective analysis disappear. However, the quality of the indicator is still dependent on two pieces of information: the respondent's age and marital status.

In general, age (or date of birth) is a problematic variable in sub-Saharan Africa. It raises data collection issues, since it is not well understood by respondents and long remained irrelevant to local practices; it also poses problems during analysis because of errors and lack of precision in reporting (Ewbank, 1981; Blanc and Rutenburg, 1990; Roger et al., 1981; van de Walle, 1968; Waltisperger, 1988). Education, migration and administrative requirements have helped to improve reporting, but data quality remains a cause for concern in most African countries (Pullum, 2006). Age inaccuracies would not have significant consequences for estimating age at marriage if they were independent of respondents' marital status. However, this is not the case. When age is unknown, the family life cycle can provide reference points for an estimate. Marital status is one of these: between two women of the same actual age, never-married women will tend to be assigned a younger age than their married counterparts, using local norms of age at marriage as a practical reference if necessary (Caldwell and Igun, 1971; Ewbank, 1981; Roger et al., 1981; Blanc and Rutenburg, 1990; Gage, 1995).

There is also a risk of error or inaccuracy in recording marital status. Where marriage formalization includes different elements, an individual could – depending on the criterion (ceremony, cohabitation, etc.) – be classified as "single" or as "married" and placed in the never-married or the ever-married

---

(3) If entry into union is early, median age at first union is correlated with the proportion of never-married women aged 15-19 (Lesthaeghe, 1989); when it is later, there is a stronger correlation with the proportion of never-married women aged 20-24. According to our calculations using the pan-African database on nuptiality (see below), the median age at first union is similarly correlated (coefficient of 0.95) with the proportion of never-married women aged 15-19 when median age at first union is under 21, and with the proportion of never-married women aged 20-24 when median age at first union is 21 or over.

category. This issue arises in particular for individuals who have uncertain or temporary marital status, or who are on the margins of the usual customs (for example, marriages in the process of being formalized, or non-cohabiting unions); in such situations, there is a risk that reporting will tend to reflect the expected response for a person of a given age or status. Some young women who have experienced a short period of marriage are likely to be recorded as "never-married" rather than as divorced or widowed; at the same time, other – somewhat older – women may be reluctant to state that they have never been married, since this is not a valued status.

In general, both censuses and surveys focus on the de facto situation, relying on the respondent's own account without laying down any precise criteria (Antoine, 2006; Lloyd, 2005). This pragmatic approach is probably a good solution: it assumes that individuals interviewed in their own homes will state spontaneously, for themselves and on behalf of the people who live with them, if they are in a union or if they have been so in the past. The imposition of specific criteria, on the other hand, might make them hesitant, introducing confusion and complexity into the recording process. In some cases, such as the last censuses in Kenya (KNBS, 2009) and Uganda (UBS, 2002), instruction manuals explicitly ask census enumerators to record reported marital status without asking for precise details. The category "married/in a union", although often handled as one questionnaire response, can also be broken down into separate items, distinguishing between consensual unions and marriages, between monogamous and polygamous situations. For example, the 1996 and 2001 South African censuses (SSA) distinguished between three types of union: cohabiting unions (living together like married partners), civil marriages and customary marriages – although the most recent census (2011) no longer differentiates between the last two categories. In Mali (INSTAT), the first three censuses (1976, 1987, 1998) categorized men's marital status by their number of wives and women's by their number of marriages; the last census (2009) uses three identical categories for both sexes: "In a monogamous marriage", "In a polygamous marriage", "Consensual union/living together". Whether systematic differences exist between surveys and censuses remains an open question: although surveys – the DHS in particular – have been more inclusive of cohabiting unions (Blanc and Rutenberg, 1990; van de Walle, 1993), censuses may have been more accurate in recording non-cohabiting marriages (Antoine, 2006 ; van de Walle, 1993). The categories that appear in questionnaires tend to reflect the variety of situations rather than a clear difference between censuses and surveys (Lloyd, 2005).

Possible errors associated with marital status, including age reporting, and the fuzzy delimitation of marital status, are constraints affecting both censuses and surveys. However, other factors are liable to generate differences in results linked to the data source. These include the conditions of data collection and the eligibility criteria.

Data collection conditions

The circumstances in which data are collected for censuses and for surveys differ in at least two ways: the personnel deployed and the status of respondents. The issue of data collection personnel – in terms of numbers, level of recruitment, training and supervision – is viewed as the weak spot of censuses, whereas surveys, in contrast, "can call on better selected, better trained and better managed personnel" (Clairin, 1988). In that regard, survey data are generally considered to be of higher quality than census data (Clairin, 1988; Tabutin, 2006). The same is true with regard to the status of respondents (Blanc and Rutenberg, 1990). In general, censuses in Africa use a household questionnaire completed with the help of one representative of the household (often the household head), whereas surveys record most information, including marital status, using an individual questionnaire completed with the person concerned.[4] Third-party responses will inevitably be less reliable. The head of a household does not necessarily know the precise marital status of each member, particularly if – as is often the case in Africa – the household includes individuals who are not his own close kin (such as a wife's relative or a young domestic servant). Moreover, the method by which census data are collected (repeating the same questions for each individual on the list of household members) does not lend itself to discussion of particular cases, and the respondent may be tempted to avoid reporting situations that are viewed as problematic, simply in order to fit into the expected categories. With an individual survey, the interviewer is actually speaking to the person who, it can be assumed, has the most precise knowledge of her own marital status. This does not prevent inaccurate reporting but does mean that it tends to be a deliberate choice on the part of the person concerned – for example, to conceal a situation associated with low social esteem. The survey protocol can be designed to limit these biases, either through the interview conditions (confidentiality, duration, attentiveness) or because a given marital status often leads to further questions and can therefore be matched against other biographical data to identify and correct inconsistencies.

Such distorting factors will affect estimates of age at marriage through their impact on measures of the proportion of never-married at the start of adult life. The conditions of census data collection increase the risk of overestimating the number of never-married women aged 15-19: the use of only one informant and the choice of recording method tend to result in young widowed or divorced women being classified as never-married, whether through ignorance or for convenience. There is some confusion around the term "single", used in questionnaires as a synonym of "never-married", since it is commonly

---

(4) Not all surveys record marital status via an individual questionnaire, however. Some (such as the first large surveys in Francophone Africa, conducted in the 1960s) consist only of a household questionnaire. Some surveys restrict the individual questionnaire to ever-married women, recording marital status through a household questionnaire, for instance the PAPFAM surveys in Arab countries. Others record marital status through both household and individual questionnaires – surveys in the MICS-2 and DHS-V programmes, for example.

taken to mean "with no partner" (including after union dissolution) – which can also contribute to this bias.[5] In some censuses – in Senegal (DPS, 2002), for example – interviewers have been alerted to the risk of error that this entails.

Eligibility and interviewer effect

Respondent eligibility criteria and the way these interact with interviewers' practices represent yet another issue that can lead to bias, but affecting this time the quality of survey data. Age transfer at the upper and lower limits of eligibility for the survey or for additional modules is a well-known phenomenon (Arnold, 1990; Rutstein and Bicego, 1990; Marckwardt and Rutstein, 1996; Pullum, 2006; Schoumaker, 2009). For example, when the individual questionnaire concerns the 15-49 age group, an imbalance between numbers aged 14 and aged 15, and between numbers aged 49 and aged 50 is often observed. Again, when a specific survey module (relating, for example, to vaccination, breast-feeding or post-partum behaviour) applies only to children aged under 5, births dated 6 years before the survey are commonly over-represented. This phenomenon can be attributed to the convergence of two factors: respondents' ignorance of their age and the tendency among some interviewers to lighten their workload by classifying individuals on the margins of the eligibility criteria as "out of range". According to the standard protocol of major surveys such as the DHS and the MICS, all women aged 15-49 recorded in the sample households should be surveyed individually. Even though her individual questionnaire will be shorter, leaving out an eligible young woman necessarily reduces the interviewer's workload.[6] According to published DHS reports, the fact that adolescent girls move around and would be obliged to return to the household to be questioned, and the possibility that interviewers are too embarrassed to ask young girls (who may not yet be sexually active) about their sexuality, both contribute to this tendency to transfer young women below the eligible age of 15 (Rutstein and Bicego, 1990; Marckwardt and Rutstein, 1996). Distortions attributed to interviewers are recognized as a classic problem in the DHS and have now become the focus of particular attention in field manuals and when re-interviewing for quality control purposes (ICF Macro, 2009, 2011). This risk of bias is smaller in censuses, because most of the questions apply to the entire population regardless of age, so there is no incentive for the enumerator to underestimate an individual's age (Rutstein and Bicego, 1990).[7] If there is age underestimation of women aged 15-16, this

---

(5) This confusion is also found in French, around the terms "*célibataire*" and "*jamais marié(e)*".

(6) The interviewer effect has also been noted in relation to another eligibility criterion: the requirement of having spent the night preceding the survey in the household. This criterion figured in the first DHS surveys but was subsequently withdrawn as it led to artificial underestimation of the number of women aged 15-49 to be interviewed (Rutstein and Bicego, 1990; Marckwardt and Rutstein, 1996).

(7) A comparison between censuses and DHS surveys dating from the late 1980s shows that this pattern of distortion is not apparent in censuses (Rutstein and Bicego, 1990).

can be assumed to mainly concern girls with none of the markers of adulthood, especially never-married women, leading to an underestimation of the never-married proportion in the 15-19 age group, and consequently of median age at first marriage in surveys.

To sum up, there are several reasons why cross-sectional measures of age at marriage may differ depending on whether census data or survey data are used. There is most probably a shared tendency to underestimate the actual value of the indicator, because lower ages are frequently attributed to young never-married women in contexts where their ages are unknown. Two additional factors come into play, combining to produce an estimate of the never-married proportion in the 15-19 age group – and hence of median age at first marriage– that is higher with census data than with survey data. These are data collection conditions on the one hand, which favour survey-based estimates, and eligibility criteria and interviewer effects, on the other, which favour census-based estimates.

To what extent are these predicted differences borne out by reality? A comparison of indicators based on data from censuses and surveys carried out in Africa over the last 50 years should shed light on the question.

II. Comparing census-based and survey-based estimates

To assess the consistency of cross-sectional estimates of age at marriage drawn from censuses and from surveys, we follow a two-stage analysis. We first calculate, for each country and each data collection operation in turn, the difference between estimates made on the basis of the two series. We then compare nuptiality trends in the different countries according to the two sources.

Data and indicator

Here we make use of tables on the distribution of population by marital status, sex and age group, brought together in INED's pan-African database on nuptiality (Hertrich, 2007a). This database includes 453 data collection operations for 1950-2010[8] from the 55 countries of Africa, 41% of which are censuses (186) and 59% national surveys (130 DHS, 46 MICS and 91 others). We have seven or more data collection operations for 70% of the countries, and at least 10 operations for 35% of the countries. The majority of countries have at least one census or survey for the period 2000-2010 (98% of countries) and at least one from the preceding decade (89% of countries). Coverage of earlier periods is poorer, but still substantial: for two thirds of countries, we have at least one pre-1970 operation.

_____

(8) Around 90% of censuses and demographic surveys carried out over this period.

The indicator used for each data collection operation is the median age of women's first union,[9] calculated from the proportions of never-married women by five-year age group.

There are two series of results for each country, the census series and the survey series. We compared them for the period covered by the two series, linking two estimates to each data collection operation: median age at marriage drawn from the census or survey in question and median age at marriage calculated by linear interpolation of the series from the other data source. The difference between these two estimates – made for the same date – defines our consistency indicator. This indicator is available for 250 operations[10] in 46 countries. In order to compare trends associated with the two types of source, an additional restriction was applied: four countries were excluded because their two series had only one point of comparison; this left 42 countries in the analysis.

Comparing estimates for all the countries of Africa

Table 1 presents an overall assessment, at the scale of the African continent and its regions, of the comparability of estimates of median age at marriage drawn from censuses and from surveys. Consistency to within ±0.5 years can be observed in half of all cases and to within ±1 year in three quarters of cases. The mean difference is 0.3 years (0.8 years on the absolute value of the differences). There is no trend towards reduction in discordances: on the contrary, the highest rates of discordance were recorded over the last decade (40% of differences of more than one year in the period 2000-2010 – double the rate for earlier periods) (Table 1). This can probably be attributed to a number of different factors, including the widespread rise in age at first union (Lloyd, 2005; Hertrich, 2007a; Shapiro, 2014; Ortega, 2014) which increases the likelihood of observing large age differences, greater diversity in pathways of entry into union which make marital status even more imprecise, and persistent dating problems (Pullum, 2006). The degree of consistency varies by region, with the best score observed in Eastern Africa (90% are consistent to within ±1 year) and the worst in Southern and Middle Africa (36%). Western and Northern Africa have intermediate scores (75% are consistent to ±1 year) (Table 1).

---

(9) i.e. the age at which the never-married proportion is 50%. This median age is calculated by linear interpolation between the age groups on either side, according to the formula:

$$Me = (x + 2{,}5) + \frac{C_{(x,x+5)} - 50}{C_{(x,x+5)} - C_{(x+5,x+9)}} \times 5,$$

where $C(x,x+5)$ is the never-married proportion in the relevant age group $(x,x+5)$.

(10) This includes each data collection operation for which an estimate associated with the other source can be calculated – a census taking place between two surveys or a survey conducted between two censuses. Operations at the extreme ends of their series are therefore not taken into account, and nor are operations falling into series that do not intersect within the time period concerned.

Table 1. Differences (census estimate – survey estimate) between estimates of median age at marriage calculated from censuses and from surveys, 1950-2010

	Distribution (%) of differences				% differences between (-1) and (+1)	Mean		Correlation coefficient between the two series	Number of operations
	< (–0.5)	(–0.5) to (+0.5)	> (+0.5)	Total		of differences	of the absolute value of differences		
Overall	17	53	30	100	74	0.3	0.8	0.90	250
By region									
Eastern	9	75	16	100	90	0.2	0.4	0.94	77
Western	19	47	34	100	75	0.3	0.7	0.90	91
Southern and Middle**	3	18	79	100	36	1.4	2.0	0.82*	39
Northern	42	56	2	100	77	– 0.5	0.7	0.96	43
Sub-Saharan Africa	12	52	36	100	73	0.5	0.9	0.88*	207
By period									
Pre- 1980	13	53	34	100	81	0.2	0.6	0.92	32
1980-89	15	67	18	100	80	0.5	0.8	0.81	55
1990-99	20	53	27	100	78	0.1	0.8	0.89	85
2000-2010	18	42	40	100	62	0.4	0.9	0.94	78

* The correlation is affected by two outliers from Botswana. When these two surveys are excluded, the correlation coefficient becomes 0.96 for the Southern and Middle region, 0.95 for sub-Saharan Africa and 0.94 for Africa as a whole.
** Given the small number of observations available for Southern Africa (5 countries, 24 observations) and Middle Africa (9 countries, 15 observations) and the similarity of the patterns observed, we have grouped these two regions together.
Source: Pan-African database on nuptiality, INED.

Comparing trends

How do these discordances show up in the national series? Are there random differences? Or do patterns of error recur, suggesting that statistical information produced by censuses is fundamentally different from that of surveys? In order to answer these questions, we grouped the countries into five categories. Series were deemed to be strictly consistent if they matched (to within approximately 0.5 years at every point) or broadly consistent if divergence was exceptional and small. Series were defined as inconsistent if differences of more than 0.5 years affected over 20% of observations; a distinction was made between the pattern where census estimates of median age at marriage were higher than survey estimates, the pattern where they were lower and the pattern where the difference varied (Appendix Table A.1, Figure 1).[11]

(11) Contrary to what might have been expected, recorded inconsistencies do not increase with the number of points of comparison. Just the opposite: the proportion of countries with consistent patterns increases with the number of data collection operations – 19% for countries with fewer than five points of comparison, 27% for those with 5-8 points and 45% for those with at least eight.

Figure 1. Trends in median age at first marriage: consistency of time series
calculated from census data and from survey data
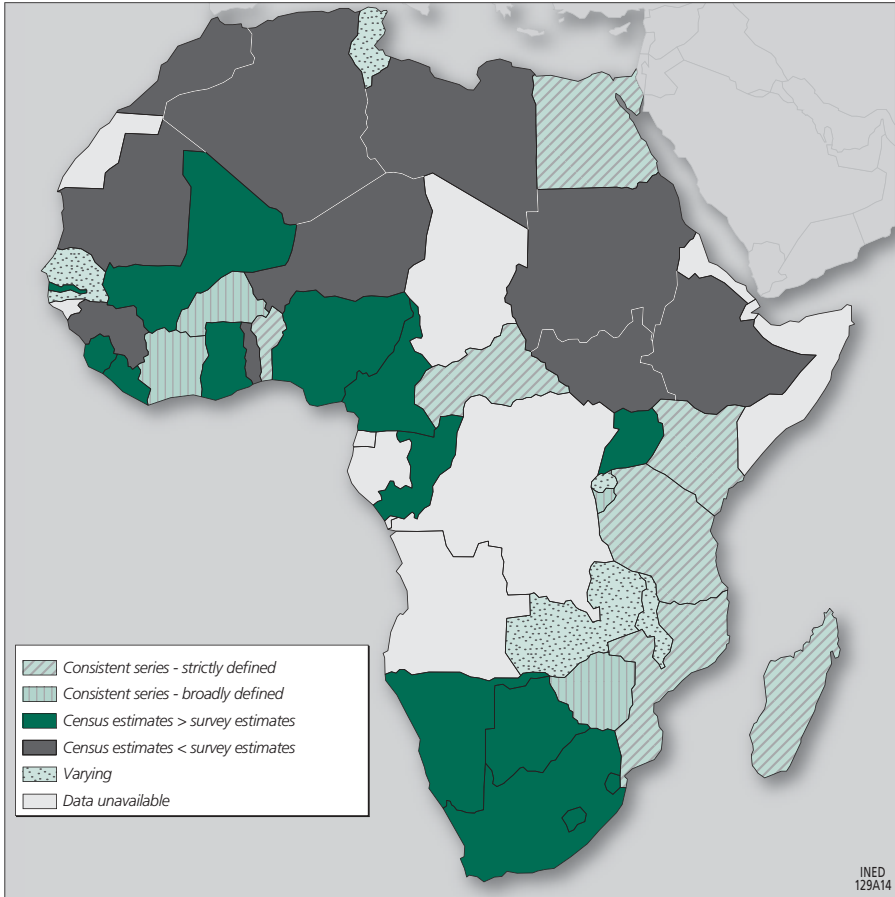

Source: Appendix Table A.1.

Only a quarter of countries (11 out of 42) displayed consistent patterns:
examples include Kenya, with near-perfect correspondence in its data series
(nine points), and Burkina Faso, with two isolated distortions out of
11 observations (Figure 2).

Regional patterns emerge for groups of countries that display inconsistencies
(Figure 1), with a contrast between the northern arc of the Arab countries,
where age at marriage estimated from surveys is almost systematically higher
than estimates from censuses, and the sub-Saharan countries, where the
discordance is most often in the opposite direction. The latter group includes
not only the countries of Southern Africa, where entry into union tends to be
relatively late, but also those of Western Africa, where it is much earlier. There
are only a few countries (five instances) that do not display a standard pattern
of discordances, and they are not in any particular region. In order to illustrate

Figure 2. Trends in median age at first marriage calculated from census data and from survey data. Examples, by type of inconsistency
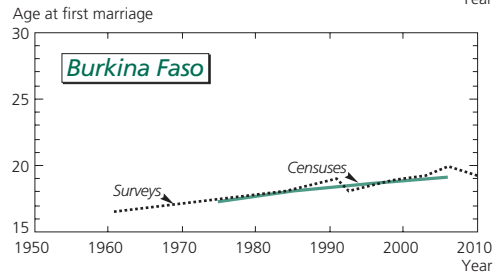
Consistent series – strictly defined

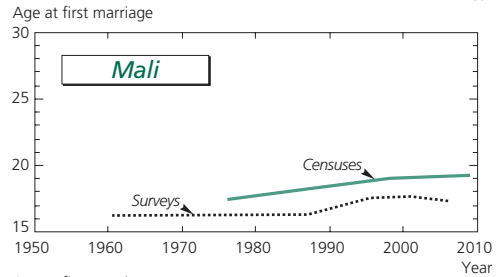*The difference between the two estimates never exceeds 0.5 years*
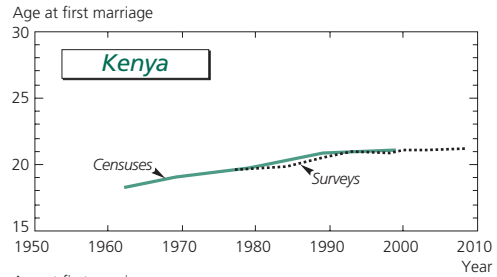
Consistent series – broadly defined

No more than 20% of data points with a difference of 0.5-1 year

Series with inconsistencies

Median age (censuses) > Median age (surveys) no more than one data point with a difference in the opposite direction
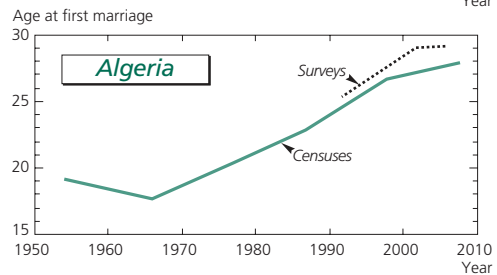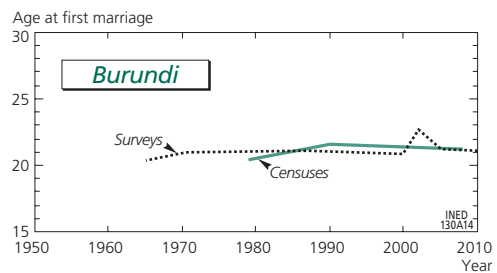
Series with inconsistencies

Median age (censuses) < Median age (surveys) no more than one data point with a difference in the opposite direction

Series with inconsistencies

Varying differences

Source: Pan-African database on nuptiality, INED.

the three distinct patterns of inconsistencies, Figure 2 provides the trends for Algeria, Mali and Burundi.

Are there two models of distortion?

The two types of comparisons (point indicators and trends) between estimates drawn from censuses and from surveys converge, confirming the existence of characteristic differences between the two sources. We can see two models of distortion, which are geographically delineated. The first pattern is predominant – and to be expected – in sub-Saharan Africa: median ages at first marriage associated with surveys are younger than estimates from census data (36% of the points of comparison and 75% of the inconsistencies). The second follows the opposite pattern and is predominant in Arab countries (42% of the points of comparison and 95% of the inconsistencies) (Table 1).

We made these comparisons for each country in turn and for the same date, so the observed differences cannot correspond to the true situation. Rather, they must result from differences between the ways that surveys and censuses treat the same reality. The two characteristic patterns of distortion north and south of the Sahara suggest that we should focus on the aspects of data collection protocols that vary between the two regions. The question of eligibility and how it is handled by interviewers represents a key issue in this debate. Most surveys conducted in Arab countries (such as those in the PAPCHILD/PAPFAM programmes) have eligibility criteria that differ from the classic protocol (as used in the DHS, for example) in their handling of marital status. Data on marital status are collected via a household questionnaire and, among women aged 15-49, only ever-married women are surveyed individually. If, when completing the household questionnaire, interviewers faced with individuals whose status is unclear prefer to consider them as ineligible, we can expect results to vary according to the eligibility criterion. As we have already seen, the hypothesis that the numbers of never-married women aged 15-19 will be underestimated because some young never-married women have been transferred into the next youngest age group is consistent with the sub-Saharan context, where ages are not known and the eligibility threshold is set at age 15. Conversely, it is logical to not find this pattern of discordances in North Africa, where ages are better known and where the constraints of eligibility with regard to marital status tend to produce an overestimation of numbers of never-married women in surveys. Survey protocols (with differing eligibility-related biases) thus offer an explanation for the two characteristic patterns of distortion north and south of the Sahara. This does not imply that censuses provide reliable estimates; rather, it means that, because the protocol they use is similar, it does not, in principle, introduce different forms of bias from one region to another.

III. Distorting factors:
an empirical examination based on MICS-2

Identifiable patterns of discordance between median ages at marriage drawn from censuses and surveys are helpful when discussing the factors underlying them, but they do not prove these factors – or enable us to evaluate whether one data source is ultimately more reliable than the other. In order to do this with complete rigour, we should compare individual records (of age and marital status) from the two sources with exact data or, failing that, make at least a comparison between the sources on the basis of individual cross-linkage. Neither of these approaches is available to us. However, some surveys do record marital status twice, in both the household and the individual questionnaires; this allows us to identify inconsistencies at the individual level and to evaluate their impact on the indicators. Although not exactly the same, to some extent the data collection conditions in the household module of a survey look like those of a census (one respondent per household, the same list of questions about each individual, non-confidential interviews, etc.), so their effects may be similar. On the other hand, the effects of certain elements that differentiate a census from a survey are impossible to capture (selection, training and supervision of interviewers), while others can be assessed only indirectly (eligibility criteria). Despite these limitations, this exercise offers the opportunity for a straightforward, empirical approach to the inconsistencies generated at the individual level by differences in protocol. This is the angle from which we shall now analyse the MICS-2 surveys, looking first at discordances in recorded marital status, then at errors in recorded age and the resulting sample distortions. Finally, we shall attempt to decide which data source should be preferred, focusing for this purpose on sub-Saharan Africa.

MICS-2: discordances consistent with the pattern observed between censuses and surveys

The second round of UNICEF surveys (EGIM/MICS-2), conducted around the year 2000, recorded women's marital status in two ways:[12] on the household questionnaire, through a question about all household members aged 15 or over; and via the individual questionnaire administered to each woman aged 15-49 (Appendix A.2). This dual record of marital status (on both household and individual questionnaires) is available for 15 sub-Saharan countries;[13] we downloaded the country databases and processed the data directly.

---

(12) In both cases, marital status was the subject of a brief, isolated question, and no other question on marital life history was included.

(13) MICS-2 surveys were conducted in eight other countries, but in seven of them marital status was recorded only on the household questionnaire. In the eighth case (Madagascar) there was complete equivalence between the marital status variables shown in the two databases. These eight countries are therefore not included in our analyses.

To accurately assess their consistency, the two declarations of marital status should be recorded independently. But this is not guaranteed by the protocol. On the contrary, the individual questionnaire design makes it possible for the interviewer to take into account and to check the marital status recorded on the household questionnaire (Appendix A.2). However, the extent of the discordances (Table 2) suggests that no practical steps were taken on any large scale to ensure consistency between the two records (except in Madagascar, which was excluded from our analyses). Since there is no guarantee of independence, and because in some instances both questionnaires might have been completed with the same person, the observed levels of discordance should be viewed as a low minimum. We used these linked data to document the mechanisms of distortion, attributing no more than a relative value to the frequency of discordances.

In the MICS data, we can again observe the same pattern of discordances as noted between censuses and surveys across the whole of sub-Saharan Africa. In the 15 countries whose data we analysed, the never-married proportion in the 15-19 age group is always higher when taken from the household questionnaire than from the individual questionnaire. The difference between the two estimates of median age at marriage is more than 0.5 years in nine of the 15 countries (Table in Appendix A.3), meaning that this situation is even more frequent than the divergence recorded for sub-Saharan Africa as a whole (36% of cases: Table 1).

Reported marital status: what about the "false never-married"?

As the individual survey of women involves a personal interview carried out, in principle, in conditions of confidentiality, we would expect it to provide better-quality information about marital status than the household survey, especially for women whose marital status is atypical or temporary. The data confirm this difference: young women who reported being "divorced" or "widowed" when interviewed individually had very often been recorded on the household questionnaire as "never-married". In 11 of the 15 countries studied, this situation was observed for over 60% of divorced or widowed women aged 15-19 (Table 2); and in five countries, this type of error was more or less systematic (90% or higher).[14] The incorrect classification of widowed and divorced women by the household survey is a factor that, alone, is enough to explain the differences in never-married proportions recorded through the household and the individual questionnaires. By comparison, consistency within the categories of "married" and "never-married", where the majority of women in this age group are concentrated, is generally high (90% or more).

---

(14) Even among the small number of young women recorded as widowed or divorced in the household questionnaire, inconsistency with the individual questionnaire is high, suggesting that respondents, including the women themselves, are reluctant to acknowledge this status.

Table 2. Comparison of marital status recorded on the individual women's questionnaire and on the household questionnaire, women aged 15-19 covered by both surveys

| Country | Declarations consistent (%) between the two sources* | | | | | | Marital status (% of 15-19 age group) | | | | | | Number |
| | Women's questionnaire | | | Household questionnaire | | | Never-married | | Divorced/widowed | | Married | | |
	Never-married	Married	Widow/Divorced	Never-married	Married	Widow/Divorced	Household questionnaire	Women's questionnaire	Household questionnaire	Women's questionnaire	Household questionnaire	Women's questionnaire	
Niger	99	98	80	97	99	92	55	54	3	3	42	42	1,238
DRC	99	95	63	98	95	80	76	75	3	3	22	22	2,921
Lesotho	96	87	63	96	87	63	79	79	1	1	20	20	1,407
Chad	98	97	63	95	97	94	59	57	4	6	37	37	1,243
CAR	94	75	37	73	96	41	50	39	16	18	34	43	3,718
Comoros	95	90	33	93	85	61	77	75	3	6	20	19	1,089
São Tomé	95	91	33	86	92	63	67	60	7	13	27	27	616
Togo	97	93	32	95	93	52	78	76	3	4	19	19	856
Sierra Leone	95	89	19	87	93	50	56	51	2	5	43	44	993
Swaziland	96	59	15	93	68	31	90	87	1	3	9	10	1,180
Gambia	95	95	10	91	89	50	65	62	1	7	34	32	1,174
Guinea-Bissau	94	91	8	87	85	34	69	64	2	8	29	27	1,558
Cameroon	99	90	5	92	98	33	76	70	1	4	24	26	1,129
Equatorial Guinea	98	44	3	67	92	33	83	57	1	10	16	33	906
Burundi	99	84	3	93	94	33	94	88	1	6	6	7	1,197

*The proportion of consistent declarations is measured with respect to the total number of women recorded with this marital status by the questionnaire under consideration (household or women's).
Note : Countries are listed in descending order by proportion of consistent declarations within the category of widowed or divorced women in the women's survey. Unweighted data and indicators.
Sources: MICS-2 surveys that recorded marital status on both household and individual questionnaires.

These indicators throw further doubt on the quality of information collected by the household survey and may appear to suggest that data from the individual surveys should be preferred. However, we must not forget another factor that works in the opposite direction: the distortion of the sample covered by the individual survey.

Sample distortion: who are the women missed by the survey?

If the survey sample includes a lower proportion of never-married women than the general population, there will be a bias towards underestimation of age at first marriage. Two factors are likely to contribute to this type of underestimation in the 15-19 age group: (a) the selection of respondents on the basis of their marital status (if never-married women are more likely to be excluded from the survey) and (b) errors in estimating age correlated with marital status (underestimation of the age of young never-married women and, possibly, overestimation of the age of ever-married adolescents). Examining the age structure of respondents who answered the individual questionnaire, the existence of such distortions seems obvious. Figure 3 illustrates this through the cases of the Central African Republic (CAR) and Cameroon. The gap between the two curves indicates the sample loss, i.e. the eligible women (aged 15-49) who should have been surveyed individually but were not. In both cases, this sample loss is especially evident in the 15-19 age group and for never-married women (in both countries, one in five never-married women aged 15-19 were not surveyed, versus one in ten ever-married women in this age group). The classic irregularities that result from age heaping are also visible, although, remarkably, 15-year-olds escape this distortion: their numbers are much lower than those of 14-year-olds. This distortion is an additional illustration of the fact that some women who should have been included in the individual surveys were classed as ineligible – and this has a direct impact on estimates of median age at first marriage, which is heavily influenced by the never-married proportion in the 15-19 age group. This phenomenon is particularly marked in the Central African Republic, which has a substantial irregularity in the age groups on either side of the eligibility range – before age 15 and after age 50.

To gain an overall picture of the impact of these distortions at the scale of the 15 countries whose data we analysed, we calculated two series of indicators for each country (Table 3). The first series relates to the effect of selection of ever-married women aged 15-19 and includes two indicators: the proportion of women recorded in the list of household members who were not surveyed individually and, in turn, the proportion of these who were never-married. The second series estimates the sample loss that results from some women aged 15-19 being classed in the 10-14 age group. In order to do this, we compared the total number aged 15-19 with a theoretical number calculated using the

Figure 3. Age distribution of women recorded on the household questionnaire
and of women covered by the individual survey, MICS-2 (2000),
Central African Republic and Cameroon
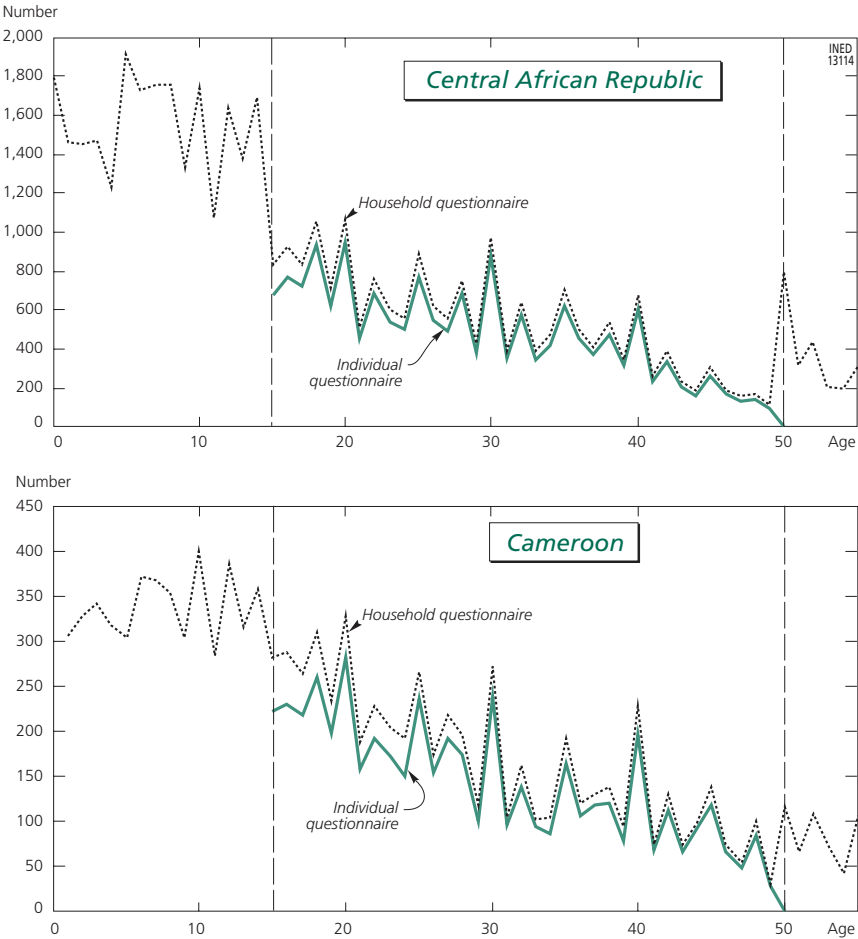
Sources: MICS-2 Surveys.

method proposed by Pullum (2006)[15] for identifying transfers from one age
group to another.

In most of the countries studied, the two types of distortion have the
cumulative effect of biasing the sample towards an underestimation of the

(15) This method (Pullum, 2006, Appendix D) looks at four age groups, hypothesizing that transfer
occurs between the two middle age groups (with numbers in the age groups on either side remaining
correct) and that there is a log-linear variation in numbers between age groups. Thus the theoretical
number aged 15-19 is calculated from the observed total numbers according to the formula:

$$Nth(15-19) = \left(No(10-14) + No(15-19)\right) \times \frac{\left(\frac{No(20-24)}{No(5-9)}\right)^{1/3}}{1 + \left(\frac{No(20-24)}{No(5-9)}\right)^{1/3}}$$

while the theoretical number aged 10-14 is: $Nth(10-14) = No(10-14) + No(15-19) - Nth(15-19)$.

Table 3. Distortion of the sample of women aged 15-19, sample loss and underestimation of women's ages

Country	Total numbers according to household survey and individual women's survey, ages 15-19			Underestimation of number aged 15-19 (%)		Number aged 15-19		
	Ratio (%)	% never-married		Household questionnaire (HQ)	Women's questionnaire (WQ)	Theoretical (estimated)	Household questionnaire	Women's questionnaire
	Number WQ/ Number HQ	Individually surveyed[1]	Not individually surveyed[2]	1 – (Number HQ/ Estimated number)[3]	1 – (Number WQ/ Estimated number)[4]	Estimated number[5]	Number	Number[6]
Niger	92	41	53	0	7	1.337	1.343	1.238
DRC	93	74	96	1	8	3.188	3.143	2.933
Burundi	86	88	99	-4	10	1.335	1.385	1.197
Chad	94	56	60	7	12	1.411	1.319	1.243
Equatorial Guinea	91	57	83	4	13	1.160	1.111	1.007
Sierra Leone	80	46	54	-6	15	1.166	1.242	993
Swaziland	91	87	86	9	17	1.453	1.319	1.199
Guinea-Bissau	80	61	73	-1	19	1.931	1.951	1.559
Cameroon	82	68	85	5	23	1.460	1.382	1.129
São Tomé	74	60	80	-2	24	902	923	683
Gambia	87	62	68	13	25	1.558	1.351	1.175
CAR	85	38	68	14	27	5.075	4.370	3.725
Lesotho	82	79	93	12	27	1.935	1.708	1.407
Togo	87	77	90	18	29	1.199	982	856
Comoros	66	76	90	-8	29	1.544	1.663	1.090

(1) Marital status as reported in the survey of individual women. The indicator is calculated on the basis of all women included in the individual survey.
(2) Marital status as reported in the household survey. The indicator is calculated on the basis of women recorded in the household database but not in the individual (women's) database.
(3) This indicator estimates the proportion of women aged 15-19 who were wrongly classed in the 10-14 age group by the household survey.
(4) This indicator measures the underestimation of the number of women aged 15-19 resulting partly from incorrect classification (transfer into the younger 10-14 age group) and partly from sample loss (eligible women not surveyed individually).
(5) Number estimated using Pullum's method (2006, Appendix D), see Footnote 15.
(6) Number of women included in the individual database.
Note: Countries in ascending order by level of underestimation of numbers aged 15-19 in the individual survey (column 4).
Sources: MICS-2 Surveys.

never-married proportion among women aged 15-19 covered by the individual women's survey. Sample loss is more than 10% in 10 of the 15 countries, and exceeds 15% in seven of these (Table 3). It is not independent of marital status: with one exception (Swaziland), the never-married proportion is always higher among young women who are recorded on the household questionnaire but not covered by the individual survey. Several factors probably contribute to this situation. Some of these undoubtedly stem – as for the DHS (Rutstein and Bicego, 1990; Marckwardt and Rutstein, 1996) – from real difficulties in reaching these girls, because of their greater mobility or their reluctance to respond in person to a survey before they have reached full adult status (shyness), or because male interviewers are embarrassed to question adolescent girls. Other factors could also depend on the interviewers, who may make more effort to interview ever-married women than never-married woman who are rarely concerned by some of the questionnaire modules (notably those dealing with children).

However, such sample loss explains only part of the shortfall of never-married women observed in the 15-19 age group. Underestimation of the age of some girls aged 15-19, leading to their classification in the 10-14 age group, is another mechanism at work. The data do not allow us to identify and analyse the particular characteristics of girls whose ages have been underestimated and who have thus been excluded from the individual survey, but it is probable that the never-married are over-represented in the 10-14 age group.

If we compare the observed total number aged 15-19 with the number that could be estimated on the basis of the age structure of the population (Table 3), we find a shortfall in the population recorded by the household survey in nine countries; this difference is reversed or absent in the other six countries. But if we take into account only women aged 15-19 surveyed individually, this measured shortfall applies to all 15 countries (Table 3): it exceeds 10% in 13 of the 15 countries and exceeds 20% in seven of them.

Household questionnaire or individual women's questionnaire: which provides a better-quality estimate?

The MICS surveys confirm the existence of two biases acting with opposite effects. On the one hand, the imprecise reporting of marital status via the household questionnaire wrongly increases the proportion recorded as never-married, while, on the other hand, distortions of the individual survey sample underestimate the numbers of never-married women aged 15-19. Can we therefore finally conclude in favour of one form of data collection over the other?

To address this question, we estimated the never-married proportion in the 15-19 age group using the data from each of the individual and household databases of the 15 MICS surveys. We proceeded as follows:

Table 4. Comparison of estimates of median age at first marriage drawn
from cross-sectional data: corrected ("probable") estimate,
estimates drawn from the household and individual women's questionnaires

Country	Median age at first marriage (AFM)			Difference: probable value – value estimated from		Distribution (%) of the difference AFM(HQ) – AFM(WQ)	
	Corrected ("Probable")	From household questionnaire (HQ)	From women's questionnaire (WQ)	Household questionnaire	Women's questionnaire	AFM(HQ) – AFM(Prob)	AFM(Prob) – AFM(WQ)
Niger	17.0	17.1	17.0	– 0.06	0.04	41	59
CAR	17.2	17.5	16.8	– 0.25	0.42	62	38
Sierra Leone	17.3	18.1	17.3	– 0.80	0.07	8	92
Chad	18.2	18.3	18.1	– 0.14	0.05	27	73
São Tomé	19.2	19.7	18.6	– 0.48	0.57	54	46
Guinea Bissau	19.3	20.0	19.1	– 0.66	0.29	31	69
Gambia	19.4	19.6	19.2	– 0.23	0.15	39	61
Equatorial Guinea	19.4	23.3	18.8	– 3.95	0.56	12	88
DRC	20.4	20.5	20.3	– 0.07	0.15	67	33
Cameroon	20.8	21.3	20.1	– 0.51	0.67	57	43
Togo	20.8	21.0	20.5	– 0.17	0.25	60	40
Comoros	21.3	21.4	20.5	– 0.02	0.79	97	3
Burundi	21.5	22.0	21.3	– 0.48	0.25	34	66
Lesotho	21.6	21.4	21.3	0.23	0.31	–	–
Swaziland	24.9	25.7	23.8	– 0.76	1.14	60	40

Note: Countries are listed in ascending order by "probable" median age at first marriage.
Sources: MICS-2 Surveys.

1) First we calculated the corrected (theoretical) numbers aged 10-14 and 15-19 using Pullum's method (2006) (see Footnote 15);

2) We then used the marital status recorded on the individual questionnaire for the group of women who took part in the survey aged 15-19;

3) For the sub-group of missing women aged 15-19 [corrected number (1) – number of who took part (2)], we assumed that the never-married proportion was the same as that observed for women aged 15-19 who appeared in the household survey[16] but not in the individual survey;

4) Finally, we recalculated the never-married proportion in the 15-19 age group as the weighted mean of the corrected estimates for the two categories, "individually surveyed" (2) and "missing" (3).

Median age at first marriage was then calculated from this proportion of never-married women in the 15-19 age group, and from the proportion given by the individual survey for the 20-24 age group. Table 4 compares this estimate

---

(16) Although this is higher than the proportion recorded for women who completed an individual survey, it is certainly lower than the true proportion.

with those provided directly by the household survey and the survey of individual women.

In the majority of countries, we note a difference between the "probable" median age and the estimate based on the household survey (which overestimates it), as well as between the "probable" median age and the estimate based on the individual survey (which underestimates it). This difference is the result of biases introduced by both forms of data collection. In half the countries, underestimation based on individual surveys has a greater impact, and in the other half, overestimation by household surveys.

We cannot reasonably conclude, therefore, that one data source is more reliable than the other, but only that neither type of source should be neglected.

Conclusion

Our analyses consistently reveal standard patterns of difference between cross-sectional indicators of nuptiality drawn from censuses and from surveys in sub-Saharan Africa. It is not so much the frequency of discordances that is striking (in three quarters of cases, estimates of median age at marriage differ by less than a year), but rather their direction: age at marriage based on census data is generally higher than age at marriage based on survey data. In the countries of sub-Saharan Africa, this pattern is observed for three-quarters of the points of discordance and for 70% of countries with discordant series of indicators.

The patterns of discordance at the scale of the continent, like those we found through a more in-depth analysis of 15 MICS surveys, clearly show that although the two sources are subject to different mechanisms of error, their results are biased in the same direction. When marital status is recorded via a household questionnaire like those used in censuses, errors regarding the marital status of young women (most particularly, categorizing widows and divorced women with never-married women) lead to an overestimation of the never-married proportion and therefore of median age at first marriage. When data on a limited section of the population is gathered, as in surveys, the number of eligible individuals is under-recorded; this results in sample deformation, creating bias in estimates of age at marriage. In sub-Saharan Africa, where the classic survey approach is to focus on women aged 15-49, overestimation of ever-married women is observed in the 15-19 age group. This is caused by two mechanisms: first, a tendency to underestimate the age of never-married young women, combined with an over-classification of girls aged under 15 as never-married, and second, lower survey coverage of never-married than of ever-married women in the 15-19 age group. Bias linked to eligibility has opposite effects for Arab countries: since never-married women are excluded from individual surveys, distortion logically works in the direction

of overestimating the numbers of never-married women – and therefore overestimating median age at first marriage.

Our analyses draw attention to data collection circumstances and their influence on data quality. We could reasonably have expected that experience gained from five decades of field operations might have gradually eliminated the problem.[17] However, it is exactly the opposite: the largest inconsistencies between estimates, in terms of both frequency and amplitude, are found during the last decade (the 2000s). In agreement with other recent studies (Bignami-van Assche et al., 2003; Johnson et al., 2009; Randall et al., 2013), our analyses call for closer consideration, when interpreting results, of interviewer effects, survey protocol design and fieldwork supervision. The distortions we observed reflect, at least in part, a certain degree of standardization of records by interviewers faced with the constraints of fieldwork and the demands of data collection. When the survey protocol requires information (such as age) that the respondent is not able to provide, interviewers are obliged to improvise – with the attendant risks that they will provide answers that appear to fit the picture (equating age with marital status, for example) and prefer responses that simplify or lighten their workload. In that regard, patterns of discordances also reflect the amount of freedom left to interviewers – suggesting that there is scope for improvement in their training and supervision. Developments in these areas are reflected in the most recent guidance on DHS field staff training (ICF-Macro, 2009, 2011), confirming that there is growing awareness of the problem.

Our results do not suggest that one data source should be preferred over the other. But they do challenge the widespread belief that surveys are of better quality than censuses. Survey data display significant and sometimes dramatic distortions, which can be easily explained by the eligibility criteria. We should therefore regard survey results with caution, in the awareness that their tendency to underestimate age at marriage inclines towards a "conservative" reading of trends which underestimates the decline in early marriages. Censuses, which tend rather to overestimate age at marriage, also have the disadvantage of being less numerous. This analysis encourages us not to give preference to one type of source, but rather to make use of both. The discordances between their results ultimately give us a more nuanced and more reliable view of trends in age at marriage than we could gain from just one data series or from simply comparing two different points in time. They do not prevent us from detecting trends; in fact, they enable us to base them on a larger number of points of comparison and to confirm them through a convergence that is strong enough

---

(17) The methodological questions that were prominent in the demographic literature on Africa in the 1960s (for example, in the studies of Francophone Africa by the Groupe de démographie africaine [African Demography Group, GDA], or in Ewbank, 1981) became less central in the decades which followed. They are now starting to be taken up again to investigate inconsistencies in some research results.

to be audible above the "noise" produced by the data's diversity and lack of precision.

APPENDICES

Table A.1. Classification of countries by consistency of census-based and survey-based series of women's median age at first marriage based on period data

A. Consistent series (11 countries)	
a. Difference between the two estimates never exceeds 0.5 years (7 countries)	b. No more than 20% of data points with a difference of 0.5 to 1 year (4 countries)
Western Africa	
• Benin	• Burkina Faso • Côte d'Ivoire
Middle Africa	
• Central African Republic	–
Eastern Africa	
• Kenya • Madagascar • Mozambique • Tanzania	• Rwanda • Zimbabwe
Southern Africa	
–	–
Northern Africa	
• Egypt	–

B. Series with inconsistencies: more than 20% of data points with a difference greater than 0.5 years (31 countries)		
a. Direction of differences:[1] the census-based median age is higher than survey-based estimate (17 countries)	b. Direction of differences:[1] census-based median age is lower than survey-based estimat (9 countries)	c. Varying direction of differences (5 countries)
Western Africa		
• Cape Verde • Gambia • Ghana • Liberia • Mali • Nigeria • Sierra Leone	• Guinea • Mauritania • Niger • Togo	• Senegal
Middle Africa		
• Cameroon • Congo • São Tomé	–	–
Eastern Africa		
• Mauritius • Uganda	• Ethiopia	• Burundi • Malawi • Zambia
Southern Africa		
• South Africa • Botswana • Lesotho • Namibia • Swaziland	–	–
Northern Africa		
–	• Algeria • Libya • Morocco • Sudan	• Tunisia
Coverage: Countries with at least two available points of comparison for the period 1950-2010. (1) No more than one data point in the series with a difference in the opposite direction. *Source:* Pan-African database on nuptiality, INED.		

Appendix A.2. Questions on marital status in MICS-2 surveys

Household Questionnaire

Excerpt from the questionnaire

9.   What is the marital status of (*name*)? **

      1 CURRENTLY MARRIED/IN UNION
      2 WIDOWED
      3 DIVORCED
      4 SEPARATED
      5 NEVER MARRIED

Excerpt from the Interviewer's Instruction Manual

Q. 9   **Marital status:** *For household members over age 15, circle the code for the response. Through-out this questionnaire 'marriage' refers to both formal and informal unions, such as living together.*

Questionnaire for Individual Women

The first question in the Contraceptive Use Module relates to marital status. The questionnaire did not include any further questions on marital life history (e.g. the first marriage).

Excerpt from the questionnaire

1. ARE YOU CURRENTLY MARRIED OR LIVING WITH A MAN?	Yes..1 No, widowed, divorced, separated............2 No, never married................................3

Excerpt from the Interviewer's Instruction Manual

Q. 1   *Check marital status in Household Listing Form (Q. 9), or ask:*
Are you currently married or living with a man?
Record the woman's status at the time of the interview. If the woman is currently married or living in an informal union, circle 1. If she is not in a union now, probe to find out if she has ever been married (circle 3 if never married) and is now widowed, divorced or separated. If one of the latter, circle 2. Remember that in this questionnaire 'married' means living in both formal and informal unions. If she is currently married (or in union) go on to Q. 2. If not, skip to next module after drawing a line through this one.
SURVEY COORDINATORS MAY DECIDE TO ALLOW SEPARATE CODES FOR 'WIDOWED', 'DIVORCED' AND 'SEPARATED'.

The question on the household questionnaire and the one on the women's questionnaire are not considered independently from each other. See information in the Interviewer's Instruction Manual:

- •"Marital status was obtained in the Household Listing Form (Q. 9): the question is repeated here as a check." (p. A1.19),
- •"Check marital status in Household Listing Form (Q. 9), or ask: Are you currently married or living with a man?" (p. A1.20)

Source: UNICEF, 2000.

Table A.3. Proportion of never-married women and median age at first marriage (period measure) based on data from household questionnaires and questionnaires for individual women

Country	% never-married aged 15-19[1]		% never-married aged 20-24[1]		Median age at first marriage[2]			Numbers[3]	
	Household questionnaire	Women's questionnaire	Household questionnaire	Women's questionnaire	Household questionnaire	Women's questionnaire	Difference	Age 15-19	Age 20-24
Niger	43	41	11	10	17.1	17.0	0.1	1,343	1,088
CAR	50	38	14	9	17.5	16.8	0.7	4,370	3,506
Sierra Leone	54	46	22	18	18.1	17.3	0.9	1,242	953
Chad	58	56	10	8	18.3	18.1	0.2	1,319	1,045
Gambia	66	62	27	27	19.6	19.2	0.4	1,351	1,308
São Tomé	71	60	23	17	19.7	18.6	1.1	923	783
Guinea-Bissau	68	61	32	26	20.0	19.1	0.9	1,951	1,780
DRC	77	74	32	31	20.5	20.3	0.2	3,143	2,390
Togo	80	77	37	33	21.0	20.5	0.4	982	890
Cameroon	75	68	42	34	21.3	20.1	1.2	1,382	1,142
Comoros	81	76	41	33	21.4	20.5	0.8	1,663	1,240
Lesotho	81	79	41	41	21.4	21.3	0.1	1,708	1,522
Burundi	94	88	45	37	22.0	21.3	0.7	1,385	833
Equatorial Guinea	83	57	55	30	23.3	18.8	4.5	1,111	798
Swaziland	89	87	66	57	25.7	23.8	1.9	1,319	1,197

(1) After weighting.
(2) Median age estimated on the basis of the never-married proportion in each five-year age group.
(3) Unweighted numbers according to the household survey.
Note: Countries listed in ascending order by median age at first marriage according to household questionnaires.
Sources: MICS-2 Surveys. De jure population. Statistical processing by the authors.

References

Antoine, Philippe (2006), "The complexities of nuptiality: from early female union to male polygamy in Africa", in G. Caselli, J. Vallin, and G. Wunsch (eds.), *Demography: Analysis and Synthesis. A Treatise in Population, Volume II*, (pp. 355-372), Boston, Elsevier Academic Press.

Antoine Philippe, Adjamagbo Agnès, Amétépé F., Béguy D., Dial Fatou, Ferrand Michèle, Tichit Christine, 2009, "Union and separation: love stories are never simple", in Antoine Philippe, Lelièvre Eva (eds.), *Fuzzy States and Complex Trajectories: Observation, Modelization and Interpretation of Life Histories*, Paris, INED; CEPED, Méthodes et Savoirs, 6, pp. 55-76.

Arnold Fred, 1990, "Assessment of the quality of birth history data in the Demographic and Health Surveys", in Institute for Resource Development, *An Assessment of DHS-I Data quality*, DHS Methodological Report no. 1, pp. 83-111.

Bignami-Van Assche Simona, Reniers Georges, Weinreb, Alexander A., 2003, "An assessment of the KDICP and MDICP data quality: Interviewer effects, question reliability and sample attrition", *Demographic Research, Special Collection* 1(2), pp. 31-76.

Blanc Ann, Rutenberg Naomi, 1990, "Assessment of the quality of data on age at first sexual intercourse, age at first marriage, and age at first birth in the demographic and health surveys", in Institute for Resource Development, *An Assessment of DHS-I Data Quality*, Maryland, Macro Systems Inc., DHS Methodological Report, no. 1.

Caldwell John C., Igun Adenola A., 1971, "An experiment with census-type age enumeration in Nigeria", *Population Studies*, 25(2), pp. 287-302.

Chae Sophia, 2011, "Forgotten marriages? Measuring the reliability of retrospective marriage histories", Communication presented at the Sixth African Population Conference organized by UAPS, Ouagadougou, 5-9 December 2011.

Clairin Rémy, 1988, "Types d'opérations de collecte", in Lohle-Tart Louis, Clairin Rémy (eds.), *De l'homme au chiffre. Réflexions sur l'observation démographique en Afrique*, Paris, CEPED-UIESP-IFORD, pp. 59-76.

DPS (Direction de la prévision et de la statistique, Sénégal), 2002, *Manuel de l'agent recenseur du Sénégal*, *Troisième recensement général de la population et de l'habitat,* République du Sénégal, Ministère de l'économie et des finances, 50 p.

Ewbank Douglas, 1981, *Age misreporting and age-selective underenumeration: Sources, patterns, and consequences for demographic analysis,* Washington D.C., National Academy Press, Committee on Population and Demography, Report n° 4, 112 p.

Gage Anastasia, 1995, *An Assessment of the Quality of Data on Age at First Union, First Birth and First Sexual Intercourse for Phase II of the Demographic and Health Surveys Program*, Calverton, Maryland, ORC Macro, Demographic and Health Surveys occasional paper 4, 53 p.

Garenne Michel, 2004, "Age at Marriage and Modernisation in Sub-Saharan Africa", *Southern African Journal of Demography*, 9(2), 2004, pp. 59-79.

Garenne Michel, 2014, "Trends in Marriage and Contraception in Sub-Saharan Africa: A Longitudinal Perspective on Factors of Fertility Decline", *DHS Analytical Studies* no. 42, Rockville, Maryland, USA.

Gᴇɴᴅʀᴇᴀᴜ Francis, Gᴜʙʀʏ Françoise, 2009, "L'observation démographique en Afrique. Leçons du passé, perspectives d'avenir, préservation et valorisation des opérations", in Marcoux Richard (ed.), *Mémoires et démographie : Regards croisés au Sud et au Nord*, Québec, AUF Presses de l'Université Laval, Cahiers du Cieq, pp. 61-97.

Gᴜʙʀʏ Fʀᴀɴçᴏɪsᴇ, 1984, *Analyse de la nuptialité*, Yaoundé, IFORD, Les documents pédagogiques, 159 p.

Hᴀᴊɴᴀʟ John, 1953, "Age at marriage and proportions marrying", *Population Studies*, 7(2), pp. 111-136.

Hᴇʀᴛʀɪᴄʜ Véronique, 1998, "Are men's and women's answers to be equally trusted? A dual collection of birth and marriage histories in a population in Mali", *Population: An English Selection*, 10(2), pp. 303-318.

Hᴇʀᴛʀɪᴄʜ Véronique, 2007a, "Nuptialité et rapports de genre en Afrique. Tendances de l'entrée en union, 1950-1999", in Locoh Thérèse (ed.), *Genre et société en Afrique,* Paris, INED, Cahier 160, pp. 281-307.

Hᴇʀᴛʀɪᴄʜ Véronique, 2007b, "Le mariage, quelle affaire ! Encadrement social et privatisation de l'entrée en union en milieu rural malien", *Sociologie et sociétés*, 39(2), pp. 119-150.

Hᴇʀᴛʀɪᴄʜ Véronique, Lᴏᴄᴏʜ Thérèse, 1999, *Rapports de genre, formation et disso-lution des unions dans les pays en développement,* Liège, IUSSP, Gender in Population Studies series, 62 p.

Hᴇʀᴛʀɪᴄʜ Véronique, Lᴀʀᴅᴏᴜx Solène, 2009, "Histoire de la nuptialité en Afrique. Peut-on comparer différentes opérations statistiques pour retracer les tendances longues de l'âge au mariage ?", in Marcoux Richard (ed.), *Mémoires et démographie : Regards croisés au Sud et au Nord*, Québec, AUF Presses de l'Université Laval, Cahiers du Cieq, pp. 204-225,

ICF Mᴀᴄʀᴏ, 2009, *Training Field Staff for DHS Surveys*, Calverton, Maryland, 53 p.

ICF Mᴀᴄʀᴏ, 2011, *Supervisor's and Editor's Manual,* Demographic and Health Surveys*, Measure DHS Basic Documentation no. 4, Calverton, Maryland, 41 p.

INSTAT (Iɴsᴛɪᴛᴜᴛ ɴᴀᴛɪᴏɴᴀʟ ᴅᴇ ʟᴀ sᴛᴀᴛɪsᴛɪǫᴜᴇ, Mᴀʟɪ), http://instat.gov.ml/

Jᴏʜɴsᴏɴ Kiersten, Gʀᴀɴᴛ Monica, Kʜᴀɴ Shane, Mᴏᴏʀᴇ Zhuzhi, Aʀᴍsᴛʀᴏɴɢ Avril, Sᴀ Zhihong, 2009, "Fieldwork-related factors and data quality in the Demographic and Health Surveys program", *DHS Analytical Studies* no. 19, Calverton, Maryland, USA, ICF Macro.

KNBS (Kᴇɴʏᴀ Nᴀᴛɪᴏɴᴀʟ Bᴜʀᴇᴀᴜ ᴏF Sᴛᴀᴛɪsᴛɪᴄs), 2009, *Enumerator's Instructions Manual*, *2009 Kenya Population and Housing Census. 24th/25th August 2009, 89 p.

Lᴀʀᴅᴏᴜx Solène, 2009, *Le mariage au Mali. Témoignages*, Paris, L'Harmattan, 187 p.

Lᴇsᴛʜᴀᴇɢʜᴇ Ron, KᴀᴜFᴍᴀɴɴ Georgia, Mᴇᴇᴋᴇʀs Dominique, 1989, "The nuptiality regimes in Sub-Saharan Africa", in Lesthaeghe Ron (ed.) *Reproduction and Social Organization in Sub-Saharan Africa*, Berkeley/Los Angeles, University of California Press, pp. 238-337.

Lʟᴏʏᴅ Cynthia B., 2005, *Growing Up Global: The Changing Transitions to Adulthood in Developing Countries*, Washington, The National Academies Press, 720 p.

Mᴀɪʀ Lucy, 1971, *Marriage*, Harmondsworth, Penguin, 221 p.

Mᴀʀᴄᴋᴡᴀʀᴅᴛ Albert M., Rᴜᴛsᴛᴇɪɴ Shea O., 1996, *Accuracy of DHS-II Demographic Data: Gains and Losses in Comparison with Earlier Surveys*. DHS Working Papers, 19, Calverton, Maryland, Macro International Inc., 22 p.

Mᴇᴇᴋᴇʀs Dominique, 1992, "The process of marriage in African societies: A multiple indicator approach", *Population and Development Review*, 18(1), pp. 61-79.

Mᴇɴsᴄʜ Barbara S., Sɪɴɢʜ Susheela, Cᴀsᴛᴇʀʟɪɴᴇ John, 2005, "Trends in the timing of first marriage among men and women in the developing world", in Lloyd C., Behrman

J., Stromquist N. P., Cohen B. (eds.), *The Changing Transitions to Adulthood in Developing Countries: Selected Studies*, Washington, DC, National Academies Press, pp. 118-171.

MENSCH Barbara S., GRANT Monica, BLANC Ann, 2006, "The changing context of sexual initiation in sub-Saharan Africa", *Population and Development Review*, 32(4), pp. 699-727.

ORTEGA ORSONA Jose A., 2014, "A characterization of world union patterns at the national and regional level", *Population Research Policy Review*, 33(3), pp. 161-188.

PULLUM Thomas W., 2006, *An Assessment of Age and Date Reporting in the DHS Surveys, 1985-2003*, DHS Methodological Reports 5, Calverton, Maryland, Macro International Inc., 86 p.

RANDALL Sara, COAST Ernestina, COMPAORE Natacha, ANTOINE Philippe, 2013, "The power of the interviewer", *Demographic Research*, 28(27), pp. 763-792.

ROGER Gilles, WALTISPERGER Dominique, CORBILLE-GUITTON Christine, 1981, *Les structures par sexe et âge en Afrique*, Paris, Groupe de démographie africaine, IDP/INED/INSEE/MICOOP/ORSTOM, 556 p.

RUTSTEIN Shea O., BICEGO George T., 1990, "Assessment of the quality of data used to ascertain eligibility and age", in Institute for Resource Development, *An Assessment of DHS-I Data Quality*, Maryland, Macro Systems Inc., DHS Methodological Report no. 1, pp. 3-37.

SCHOUMAKER Bruno, 2009, "Stalls and reversals in fertility transitions in Sub-Saharan Africa: Real or spurious?", Louvain-la-Neuve, Université catholique de Louvain, Documents de travail du Sped no. 30.

SHAPIRO David, GEBRESELASSIE Tesfayi, 2014, "Marriage in Sub-Saharan Africa: Trends, determinants, and consequences", *Population Research Policy Review*, 33(2), pp. 229-255.

SSA (STATISTICS SOUTH AFRICA), http://www.statssa.gov.za

TABUTIN Dominique, 2006, "Information systems in demography", in Caselli Graziella, Vallin Jacques, Wunsch Guillaume (eds.), *Demography: Analysis and Synthesis,* Vol. 4. Ch. 121, Amsterdam, Elsevier, pp. 493-522.

TABUTIN Dominique, VALLIN Jacques, 1975, "La nuptialité", *Sources et analyse des données démographiques. Application à l'Afrique d'expression française et à Madagascar.* Part three, Volume II, Paris, INED/INSEE/MICOOP/ORSTOM, pp. 3-56.

TABUTIN Dominique, SCHOUMAKER Bruno, 2004, "The demography of sub-Saharan Africa from the 1950s to the 2000s. A survey of changes and a statistical assessment", *Population, English Edition,* 59(3-4), pp. 457-556.

UBS (UGANDA BUREAU OF STATISTICS), 2002, *Enumerators' Instruction Manual, Census 2002,* 46 p.

UNICEF, 2000, *Manual on Multiple Indicator Cluster Surveys: The end Decade Manual and Questionnaire*, New York, UNICEF. [Appendix A1 : Interviewer instructions] http://www.childinfo.org/mics2_manual.html

UNICEF, MULTIPLE INDICATOR CLUSTER SURVEYS / MICS2, http://www.childinfo.org/mics2.html

UNITED NATIONS, 1984. *Manual X: Indirect Techniques for Demographic Estimation*, New York, Population Division, Department of International Economic and Social Affairs, http://www.un.org/esa/population/publications/Manual_X/Manual_X.htm

UNITED NATIONS, 2008, *World Marriage Data 2008*, United Nations, Department of Economic and Social Affairs, Population Division.

UNITED NATIONS, 2013, *World Marriage Data 2012*, United Nations, Department of Economic and Social Affairs, Population Division.

van De Walle Etienne, 1968, "Marriage in African censuses and inquiries", in William Brass et al., *The Demography of Tropical Africa*, Princeton, Princeton University Press, pp.183-238.

van de Walle, Etienne, 1993, "Recent trends in marriage ages", In K.A. Foote, K.H. Hill, and L.G. Martin (eds.), *Demographic Change in Sub-Saharan Africa*, Washington, DC, National Academy Press, pp. 117-152.

Waltisperger Dominique, 1988, "Les structures par âge", in Lohle-Tart Louis, Clairin Rémy (eds.), *De l'homme au chiffre. Réflexions sur l'observation démographique en Afrique*, Paris, CEPED-UIESP-IFORD, pp. 203-216.

Westoff Charles, 2003, *Trends in Marriage and Early Childbearing in Developing Countries*, Calverton, Maryland, DHS Comparative Reports n° 5, ORC Macro, 53 p.

Véronique HERTRICH, Solène LARDOUX • ESTIMATING AGE AT FIRST UNION IN AFRICA. ARE CENSUS AND SURVEY DATA COMPARABLE?

This article considers whether survey and census data offer comparable bases for estimating trends in women's age at marriage in Africa. It uses the indicator of median age at first marriage calculated from the proportion of never-married women by age. It draws upon two bodies of data: first, a pan-African nuptiality database is used to assess differences between estimates drawn from the two types of source at the scale of the whole continent (453 censuses and national surveys undertaken since 1950 in the 55 countries of Africa); and second, data from 15 MICS surveys which record marital status twice (each respondent is included on both a household and an individual questionnaire) are analysed to pinpoint inconsistencies. The median age at first marriage is generally higher when estimated from census data than from survey data. Several error mechanisms combine to create this effect. In censuses, imprecise recording of marital status leads to overestimation of numbers never-married, and therefore to overestimation of median age at marriage. In surveys, meanwhile, the tendency to underestimate young women's age, thereby excluding a disproportionate number from the survey sample of women aged 15-49, and the less thorough coverage of never-married women lead to under-representation of those never-married and therefore to underestimation of age at marriage. This analysis does not suggest that one type of source should be preferred over the other, but rather that neither source should be neglected.

Véronique HERTRICH, Solène LARDOUX • ÂGE À L'ENTRÉE EN UNION DES FEMMES EN AFRIQUE. LES DONNÉES DES ENQUÊTES ET DES RECENSEMENTS SONT-ELLES COMPARABLES?

L'article évalue la comparabilité des enquêtes et des recensements pour estimer les tendances de l'âge au mariage des femmes en Afrique. L'indicateur utilisé est l'âge médian au premier mariage tiré de la proportion de célibataires par âge. Deux corpus de données sont utilisés : d'une part, une base panafricaine sur la nuptialité qui permet d'évaluer à l'échelle du continent les écarts entre les estimations tirées des deux types de sources (453 recensements et enquêtes nationales réalisés depuis 1950 dans les 55 pays africains); d'autre part, 15 enquêtes MICS dont le double enregistrement de la situation matrimoniale, sur les questionnaires « ménage » et « individu », permet de préciser les facteurs de discordance. L'âge médian au mariage généralement plus élevé d'après les recensements que d'après les enquêtes. Plusieurs mécanismes d'erreurs sont en jeu. Côté recensements, les imprécisions sur la situation matrimoniale conduisent à une surestimation des célibataires et donc de l'âge médian au mariage. Côté enquêtes, la tendance à sous-estimer l'âge des jeunes femmes en deçà du critère d'éligibilité de 15 ans et la moins bonne couverture des célibataires par l'enquête conduisent à une sous-représentation des célibataires et donc à une sous-estimation de l'âge au mariage. Plutôt que de privilégier l'une ou l'autre source, nos analyses encouragent à n'en négliger aucune.

Véronique HERTRICH, Solène LARDOUX • EDAD DE LAS MUJERES A LA UNIÓN EN ÁFRICA. ¿LOS DATOS DE LOS CENSOS Y LOS DE LAS ENCUESTAS SON COMPARABLES?

El artículo evalúa la comparabilidad de las encuestas y de los censos para estimar las tendencias de la edad al matrimonio de las mujeres en África. El indicador utilizado es la edad mediana al primer matrimonio calculada a partir de la proporción de solteras por edad. Dos conjuntos de datos son empleados: por un lado, una base panafricana sobre la nupcialidad que permite evaluar las diferencias entre las estimaciones sacadas de los dos tipos de fuentes (453 censos y encuestas nacionales realizadas desde 1950 en 55 países africanos); por otro lado, 15 encuestas MICS (Mulitiple Indicator Cluster Survey) cuyo doble registro del estado matrimonial en los cuestionarios "hogar" e "individuo", permite precisar los factores de discordancia. Una edad mediana al matrimonio generalmente más elevada según las encuestas que según los censos. Varios mecanismos están en juego. En los censos, las imprecisiones sobre el estado matrimonial conducen a una sobrestimación de las solteras y en consecuencia de la edad media al matrimonio. En las encuestas, la tendencia a subestimar la edad de las mujeres jóvenes por debajo del criterio de elegibilidad de 15 años así como una cobertura de las solteras menos completa por las encuestas, conducen a una subestimación de la edad al matrimonio. Más que privilegiar una u otra de las dos fuentes, nuestros análisis aconsejan no descuidar ninguna de ellas.

Translated by Karen George.