Using census data to estimate old-age mortality for developing countries

Nan Li and Patrick Gerland[1][2]

Paper prepared for Session 17-05: Indirect methods of mortality and fertility estimation: new techniques for new realities (Tuesday, August 27th 2013, 13:30pm - 15:00pm) in *XXVII IUSSP International Population Conference*. Busan, Korea.

Abstract

Thanks to substantial improvements in child and adult survival through most parts of the world in the last decades, old-age mortality accounts now for large fractions of death. Yet for many developing countries, old-age mortality is often only inferred by model life tables using mortality data at young ages, or sometimes at young and adult ages; and reliable estimates of old-age mortality using data collected from old-age population can hardly be found. Based on the fact that migration is rare and death risk is high at old ages, this paper proposes a new indirect method, namely the Census method, to estimate old-age mortality, using census data on old-age population. This new indirect Census method aims to eliminate the effects of age-reporting errors, and is composed of three models. The first model is the variable-r method that converts the census populations into the person-years of the underlying stationary population. The second is an adjustment model, which uses a common relationship between the survival ratios that is found in model life tables to eliminate the effects of age-reporting errors in censuses. And the third is the extended Gompertz model, which estimates the number of survivors at exact ages of the underlying stationary population based on the most commonly observed mortality pattern. Examples are provided using census data from developing countries in Africa and Asia.

Background

Data on death collected from census or registration are less reliable in many developing countries. This reality leads to using survey data to measure mortality levels. Based on methods such as the children ever born and survival (Brass 1964; Brass and Coale 1968; United Nations 1983), young-age (e.g., under the age of 1 or 5 years) mortality are estimated, for instance, in the World Fertility Surveys (see http://opr.princeton.edu/archive/wfs). Using these measures, and combining with reliable data from vital registration and census, the United Nations Population Division (United Nations 2013) and the United Nations Children's Fund (UNICEF, http://www.childmortality.org) has long been regularly publishing young-age (under the age of 1 and/or 5 years) mortality for virtually all countries back to

1. Nan Li and Patrick Gerland, Population Division, United Nations. Correspondence to Nan Li, 2 UN Plaza-DC2-1906, New York, NY 10017, USA; Tel: (212)963-7535; Fax: (212) 963-2638; E-mail: li32@un,org.

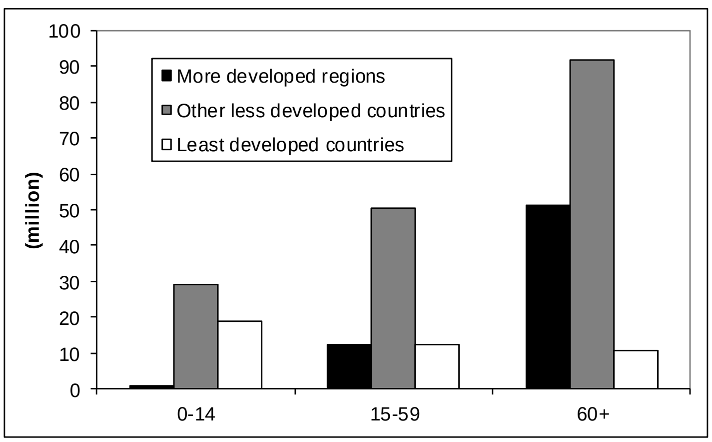
2. Views expressed in this paper are those of the authors and do not necessarily reflect those of the United Nations. The designations employed in this paper and the material presented in it do not imply the expression of any opinion whatsoever on the part of the Secretariat of the United Nations concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

1970s or 1950s. Following the estimates of young-age mortality, surveys such as the Demographic and Health Surveys (DHS, http://www.measuredhs.com) started to measure adult (e.g., between the age of 15 and 60 years) mortality for an increasing number of developing countries, based mainly on the data and method of sibling survival (United Nations 2002). Combining with census and vital registration data, Rajaratnam and colleagues (2010) estimated adult mortality for 187 countries from 1970 to 2010. Utilizing young-age mortality and one-dimension model life tables (e.g., inputting only infant mortality;Coale and Demeny (1966 and United Nations (1982), one may produce a life table, according to which mortality at any age can be inferred. Using young-age and adult mortality and two-dimension model life tables (e.g., inputting infant and adult mortality; see Murray et al. (2003;Wilmoth et al. (2011), one could do such inferences better. In fact, these measures, and together with reliable data from census and vital registration, are used by the World Health Organization (WHO, http://www.who.int/gho/countries/en/) to provided life tables for almost all the countries. These achievements significantly enriched our knowledge about the mortality of developing countries.

While child and adult mortality of most countries are systematically studied, old-age mortality has not held proper attention, although it is the elephant in the room.

It would be mistaking to consider old-age mortality is less important than at young-age and at adulthood, especially in estimating the global burden of disease (Wang et al. 2012). This is because 55% of deaths worldwide in 2005-2010 were estimated to be aged 60 years and older (United Nations 2013): 153 million out of a total of 277 million deaths occurred in 2005-2010 in that age group (Figure 1) and *more than half of all deaths occur at age 60 or higher in nearly all developing regions, except in sub-Saharan Africa where only one fourth or less happen at those ages*.

Figure 1. Number of deaths in 2005-2010 by broad age groups age and development group



Source: United Nations (2013)

Knowing young-age and adult mortality, old-age mortality could be inferred using model life tables. This is a progress from using only young-age mortality and model life tables to infer old-age mortality, but it is still not an estimation based on the data of old-age population. How to systematically estimate old-age mortality for developing countries, when the numbers of death of vital registration and census are unreliable, is perhaps still a field to be explored. In principle, old-age mortality could be estimated using indirect methods such as orphanhood or widowhood methods (Timæus 2013a, 2013b; United Nations 1983, 2002), and using survey data if proper information were collected. For example, if information on the survival status of the parents of the respondents is collected in a census or survey, then old-age mortality could be directly estimated based on these orphanhood data. Information collected by proxy respondents about their children, siblings and parents is indirect, but can efficiently enlarge the potential sample size to estimate deaths and other rare events. In practice, however, indirect information about old-age mortality has yet to be collected; and directly estimating old-age mortality using deaths and exposure survey data (Bendavid, Seligman and Kubo 2011; INDEPTH Network 2002) is still unable to provide acceptable results, due mainly to sample size limitation or sample selection issues. On the other hand, this paper offers an approach to estimate indirectly old-age mortality using data from population censuses and addressing issues related to errors of age reporting, assuming that the difference in completeness between successive censuses is negligible.

For young age-mortality, commonly used indicators are the infant mortality rate ${}_1q_0$ or the under-five mortality rate ${}_5q_0$, namely the probability of dying between ages 0 and 1 or 0 and 5, respectively. A common measure of adult mortality is the ${}_{45}q_{15}$, which is the probability of dying between ages 15 and 60, where the 45 represents the years to reach 60 from age 15. While a standard measure of old-age mortality still remains to emerge from future studies, this paper proposes to use ${}_{15}q_{60}$ as measure of old-age mortality, namely the probability of dying between ages 60 and 75, or simply the old-age mortality. The main reason to choose 60 as the lower bound is simply because it is a commonly used starting age for old-age population, and the most commonly used statutory retirement age (United Nations 2012). The reason to choose 75 as the upper bound is that most developing countries have been tabulated census data with an open age group starting at this age, especially in earlier years when populations were smaller.

In applying the variable-r method, the general obstacle is its assumption of zero migration (Bennett and Horiuchi 1981), or its requirement on the accurate numbers of net migration (Preston and Coale 1982), which is more difficult to obtain than to estimate deaths. This obstacle should become trivial at ages of 60 years and old, when there are fewer migrants. Further, the increasing deaths at old ages would make the effect of the migration negligible. In applying variable-r method to developing countries, however, the remaining difficulty is the age heaping and other errors in censuses, particularly at ages ending with 0. Based on the common age patterns of survival ratios, this paper proposes a procedure to correct the effects on old age mortality rates caused by age heaping and other reporting errors (e.g., misstatement, exaggeration) that are common in censuses of developing countries (Ewbank 1981; Preston, Elo and Stewart 1999).

Another issue is that the variable-r method produces only the person-years for the underlying stationary population, while the number of survivor at exact ages are still unknown. These numbers are hard to estimate for developing countries, given the issue of age heaping and the lack of reliable population data between censuses. This issue is addressed here using an extended Gompertz model.

Cautions must also be made, however, that the assumption of similar completeness of two successive censuses, which is required by the variable-r method, is still necessary in this paper. Nonetheless, it can be assumed that old people are less mobile, and therefore are easier to count in censuses. Hence, this necessary assumption stands better in this study than in others focusing on younger ages. Nevertheless, under the assumption that census coverage improves over time, any improved completeness for the most recent census would lead to underestimation of old age mortality. In situations when the number of death by age between the censuses could be obtained from sources independent from the census population, the difference in completeness between two successive censuses could be estimated by the General Growth Balance method (Hill 1987), and its effect on estimating old-age mortality could be adjusted. In cases when the post-enumeration surveys are available and reliable for the censuses, the effect of the differential census coverage completeness between successive censuses could also be adjusted. In addition, instead of relying only on conventional intercensal periods analyzing two successive censuses, more robust results should also be obtained over longer intercensal period using multiple censuses. Evaluation of final old-age mortality results by sex can also reveal further insights about the plausibility of the underlying assumption about differential census coverage between censuses (e.g., male mortality substantially much lower than female old-age mortality could be one of the symptoms).

The Census method

The method proposed here, namely the Census method, is composed of three parts: the variable-r method, the adjustment of errors on age, and the extended Gompertz model.

The variable-r method

The target of our estimation is the probability of dying between ages 60 and 75,

$$ {}_{15}q_{60} = 1 - \frac{l_{75}}{l_{60}}, \qquad (1)$$

where l_x represents the number of survivors at age x. Let *p(x, t)* be the observed number of population in age group *[x,x+5)* in a census at time *t*, where *x*=60, 65, 70. The growth rates at age *x* are computed as

$$ r(x) = Log[\frac{p(x,t_2)}{p(x,t_1)}]/(t_2 - t_1), \quad x = 60,65,70. \qquad (2)$$

And the accumulated growth rates are accordingly

$$ \begin{aligned} s(60) &= 2.5r(60), \\ s(65) &= 5r(60) + 2.5r(65), \\ s(70) &= 5[r(60) + r(65)] + 2.5r(70). \end{aligned} \qquad (3)$$

Further, the between-census middle-point population in age group *[x,x+5), N(x),* are estimated as

$$ N(x) = \sqrt{p(x,t_1)p(x,t_2)}, \quad x = 60,65,70. \qquad (4)$$

Using the variable-r method (see also Preston, Heuveline and Guillot 2001), the person-years lived in 5-year age group *[x,x+5)*, L_x, in the underlying stationary population, are obtained as

$$ L_x = N(x)\exp[s(x)], x = 60,65,70. \qquad (5)$$

When *r(x)* is constant over age and *N(x)* is the age-specific number of a stable population, the L_x in equation (5) is obviously the age-specific number of the underlying stationary population. When *r(x)* changes over age and *N(x)* represents a non-stable population, the variable-r method indicates that the L_x in (5) is still the age-specific number of the underlying stationary population.

The adjustments

The variable-r method assumes that the errors in enumerating population are constant with age, and identical in the two successive censuses. But in developing countries the errors often occur unevenly across age. A typical example is age heaping. When such errors are severe, the L_x resulted from variable-r method, namely the variable-r L_x, would show implausible patterns of increasing with age, which cannot occur in a stationary population. When such implausible situations occur, adjusting L_x is necessary.

It is hard to find a proper basis to adjust enumerating errors in a real population, which is affected by historical fertility, mortality and migration. Slightly different age ranges could also be considered as long as they minimize the potential influence of migrations, underreporting and age exaggeration. Murray et al. (2010) in their validation study of death registration methods recommend using the age ranges 40-70 (GGB), 55-80 (SEG) and 50-70 (GGBSEG). But a stationary population is only determined by mortality. Thus, it is possible to find a proper basis to adjust age errors for stationary populations. As is shown in the Supplementary Appendix S1, there is indeed a common relationship between the survival ratios $S_{60} = \frac{L_{65}}{L_{60}}$ and $S_{65} = \frac{L_{70}}{L_{65}}$ among model life tables, which is $S_{65} = -0.29 + 1.27 \cdot S_{60}$ based on the UN general family for females (see Supplementary Appendix S1 for further details). This relationship is called the model line, and is used as the basis to adjust age errors in this paper.

When the observed survival-ratio point, (S_{60}, S_{65}), is above the model line, or when the survival ratio is abnormally rising with age, the difference between the survival-ratio point and the model line is caused mainly by age heaping. Accordingly, the adjustment is to move the survival-ratio point into the model line, assuming that the heaping ratio at age 60 equals to that at age 70. As indicated in the Supplementary Appendix S2, this adjustment is

$$\begin{aligned} \hat{L}_{60} &= L_{60} - \frac{L_{60}}{L_{70}}\Delta, \\ \hat{L}_{65} &= L_{65} + \Delta, \\ \hat{L}_{70} &= L_{70} - \Delta, \end{aligned} \tag{6}$$

where

$$\begin{aligned} \Delta &= \frac{-B + \sqrt{B^2 - 4AC}}{2A}, \\ A &= b - a\frac{L_{60}}{L_{70}} - \frac{L_{60}}{L_{70}}, \\ B &= a(L_{60} - \frac{L_{60}}{L_{70}}L_{65}) + 2bL_{65} + L_{60} + \frac{L_{60}}{L_{70}}L_{70}, \\ C &= L_{65}(aL_{60} + bL_{65}) - L_{60}L_{70}, \\ a &= -0.29, \quad b = 1.27. \end{aligned} \tag{7}$$

On the other hand, when the survival-ratio point is below the model line, the difference between the survival-ratio point and the model line is caused by nonspecific errors. Accordingly, the adjustment is to move the survival ratio point into the model line through minimal distance. As is indicated in the Supplementary Appendix S3, the minimal adjustment is

$$\hat{S}_{60} = \frac{-ab + S_{60} + bS_{65}}{1+b^2}, \qquad \hat{S}_{65} = a + b\hat{S}_{60}. \tag{8}$$

$$\hat{L}_{60} = w\frac{L_{60} + \hat{S}_{60}L_{65} + \hat{S}_{60}\hat{S}_{65}L_{70}}{1+\hat{S}_{60}^2 + \hat{S}_{60}^2\hat{S}_{65}^2} + (1-w)L_{60},$$
$$\hat{L}_{65} = w\hat{S}_{60}\frac{L_{60} + \hat{S}_{60}L_{65} + \hat{S}_{60}\hat{S}_{65}L_{70}}{1+\hat{S}_{60}^2 + \hat{S}_{60}^2\hat{S}_{65}^2} + (1-w)L_{65}, \tag{9}$$
$$\hat{L}_{70} = w\hat{S}_{65}\hat{S}_{60}\frac{L_{60} + \hat{S}_{60}L_{65} + \hat{S}_{60}\hat{S}_{65}L_{70}}{1+\hat{S}_{60}^2 + \hat{S}_{60}^2\hat{S}_{65}^2} + (1-w)L_{70},$$

where $0 \le w \le 1$ is the weight, and is used as 0.5.

The extended Gompertz model

At the ages of interest in this paper, it is well known that the force of mortality rises roughly exponentially, which is called the Gompertz law. The force of mortality at age x is defined as

$$\mu(x) = -\frac{dl_x}{l_x dx}, \tag{10}$$

and the Gompertz model is written as

$$\mu(x) = \mu(60)\exp[g(x-60)], \quad g > 0, \tag{11}$$

where $\mu(60)$ is the force of mortality at age 60 and *g* the rate of mortality increase, they are the two parameters of the Gompertz model.

Combining (10) and (11), the Gompertz model can be extended to describe l_x:

$$l_x = l_{60}\exp[-\int_{60}^{x}\mu(y)dy] = l_{60}\exp\{-\frac{\mu(60)}{g}[\exp[g(x-60)]-1]\}. \tag{12}$$

The extended Gompertz model includes an additional parameter, l_{60}, which could be estimated together with the other two parameters, and the starting-age population is therefore solved.

There are three unknown parameters in (12), namely, l_{60}, $\mu(60)$, and *g*; and there are three equations to solve these parameters according to the definition of L_x:

$$\hat{L}_{60} = \int_{60}^{65} l_x dx = l_{60}\int_{60}^{65}\exp\{-\frac{\mu(60)}{g}[\exp[g(x-60)]-1]\}dx,$$
$$\hat{L}_{65} = l_{60}\int_{65}^{70}\exp\{-\frac{\mu(60)}{g}[\exp[g(x-60)]-1]\}dx, \tag{13}$$
$$\hat{L}_{70} = l_{60}\int_{70}^{75}\exp\{-\frac{\mu(60)}{g}[\exp[g(x-60)]-1]\}dx.$$

The solution of (13) can be found using numerical computation methods (e.g., non-linear optimization routine), which provides the estimate of old-age mortality as:

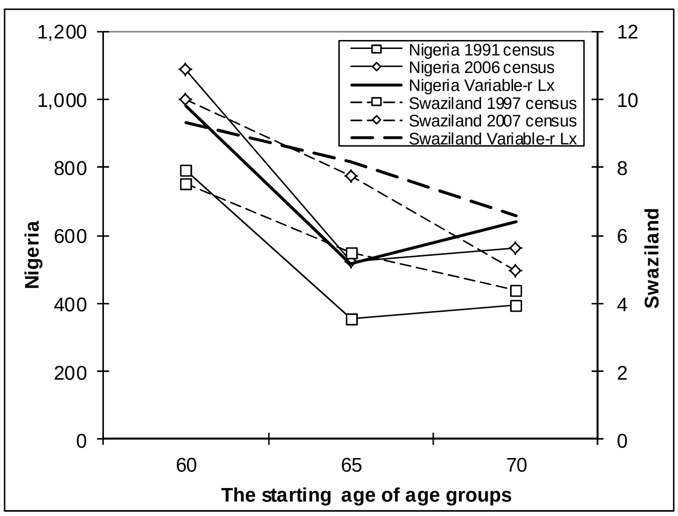
$$ {}_{15}q_{60} = 1 - \frac{l_{75}}{l_{60}} = 1 - \exp\{-\frac{\mu(60)}{g}[\exp[g(75-60)]-1]\} . \qquad (14)$$

Illustrations

To illustrate the above estimation, data of female population from the censuses of Japan in 2000 and 2005, Sweden in 1950 and 1960, Nigeria in 1991 and 2006, and Swaziland in 1997 and 2007 are chosen (all census data in this paper are collected from successive editions from the Demographic Yearbook (United Nations. Statistical Division. 1948-)). All the censuses are for the two most recent successive years except Sweden, for which the years are selected to reflect relatively lower survival ratios. Japan and Sweden are known for high quality of data, Nigeria is the most populous country of Africa, and Swaziland suffered the highest HIV adult prevalence rate in the world in 2005-2010.

The situations of two African countries' census data on female populations in the three age groups of interest are shown in Figure 2. For Nigeria, age heaping on 60 and 70 years of age was remarkable in both censuses. Applying variable-r method on such data, results are unacceptable because $L_{65} < L_{70}$. On the other hand, age heaping was not obvious in the two censuses of Swaziland, and the resulted L_x look usable. To apply variable-r method to Nigeria, adjustments are necessary.

Figure 2. Enumerated female population in 5-year age group for age 60 and over in 1991-2006 Nigeria and 1997-2007 Swaziland censuses (in thousands)

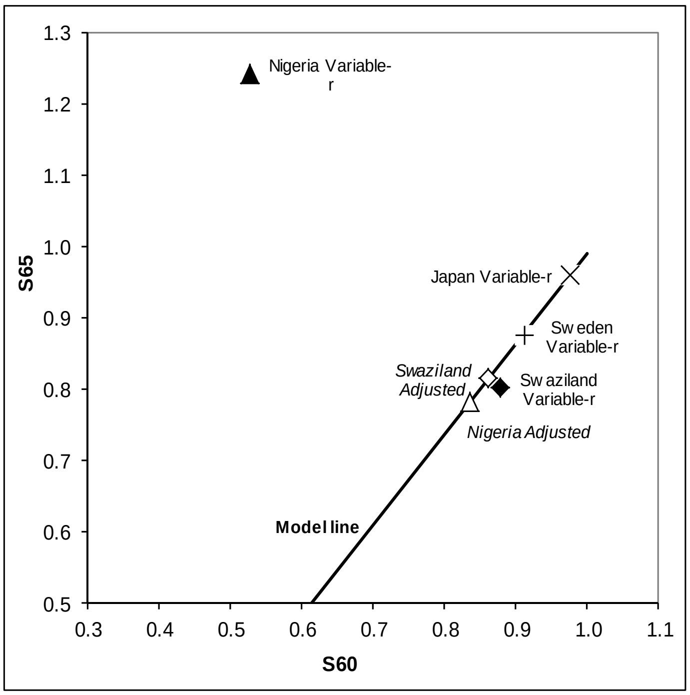


Source: computations by authors

Applying the adjustments to the data of the above four countries, results are shown in Figure 3. It can be seen that for Japan and Sweden, the survival-ratio points obtained from variable-r method are almost on the model line and therefore the adjustment makes virtually no difference. For Swaziland, point (S_{60} , S_{65}) is slightly below the model line, which could not be convincingly explained by age heaping on 65 years of age. This is caused by nonspecific reasons, and hence is adjusted minimally. The effect of the minimal adjustment is small (the weight is used as 1 for illustration purpose).

For Nigeria, point (S_{60} , S_{65}) is far above the model line because of huge age heaping on 60 and 70 years. The basis of adjusting the effect of age heaping is not minimizing the adjustment, but the assumption that the heaping ratios at ages 60 and 70 are the same. If a minimal adjustment (see Supplementary Appendix S3) was used for Nigeria, the adjusted survival-ratio point, $(\hat{S}_{60}, \hat{S}_{65})$, would be closer to that of Japan, because the segment that connects point (S_{60}, S_{65}) and ($\hat{S}_{60}, \hat{S}_{65}$) must be orthogonal to the model line through the minimal adjustment. This indicates that the minimal adjustment does not make sense for severe age-heaping situations. On the other hand, the age-heaping adjustment leads to $\hat{S}_{60}$ and $\hat{S}_{65}$ to be slightly smaller than that of Swaziland, a country with the highest AIDS prevalence. Does it make sense? It should, if the heaping ratio at age 60 is similar to that at age 70, and if the completeness of one census is similar to another.

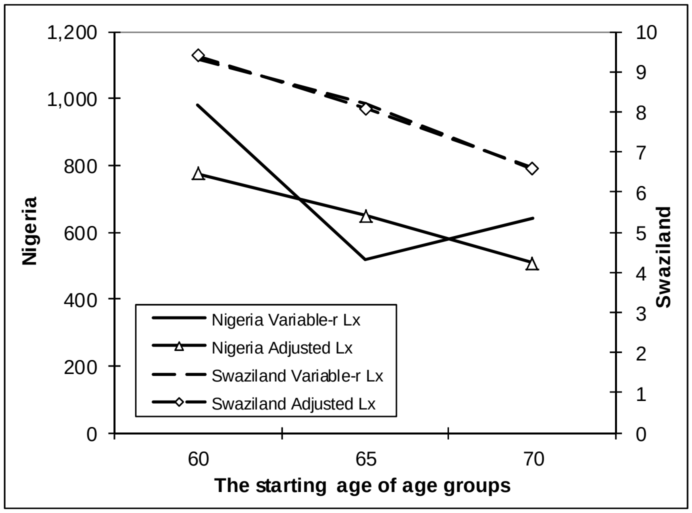
Figure 3. Variable-r and adjusted survival ratios of females for selected countries



Source: computations by authors

In Figure 4, it can be seen the effect of the age-heaping adjustment is significant for Nigeria. In fact, any adjustment that can fix the Nigerian variable-r L_x must be significant. On the other hand, the effect of the minimal adjustment is small for Swaziland, because nothing looks abnormally wrong.

Figure 4. Variable-r (un)adjusted Lx in 5-year age group for female age 60 and over in 1991-2006 Nigeria and 1997-2007 Swaziland censuses (in thousands)



Source: computations by authors

Applying the extended Gompertz model to the adjusted L_x, the estimates of old-age mortality are shown in Table 1. It can be seen that for Nigeria, the ${}_{15}q_{60}$ of males is smaller than for females, which is unusual among developed countries. But it may not be unusual for least developed countries. For example, the DHS (2008) reported higher death rates for females than for males at ages 15-40, for Nigeria. On the other hand, this unusual situation could also be a result of difference in completeness between the two censuses (e.g., greater improved completeness for males than for females.

Table 1. Estimates of old-age mortality ${}_{15}q_{60}$

	Period	Males	Females
Japan	2000-2005	0.214	0.099
Sweden	1960-1965	0.370	0.292
Nigeria	1991-2006	0.356	0.479
Swaziland	1997-2007	0.624	0.419

Applications

Applications to African countries

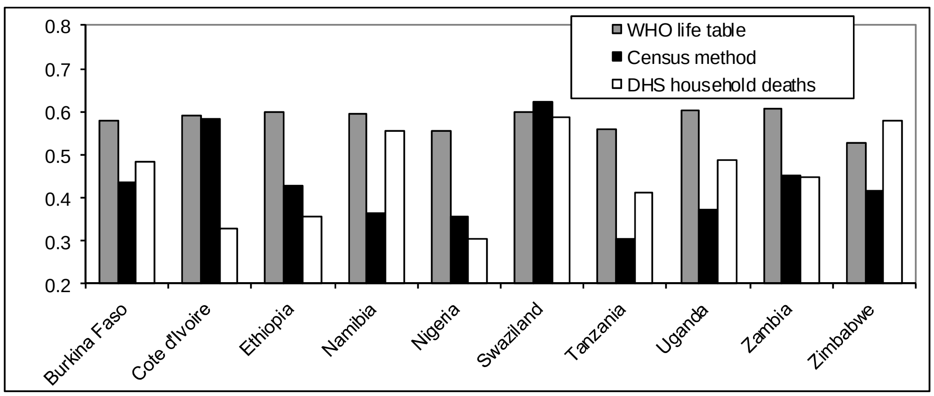
In this section, ten African countries for which DHS collected household deaths in the 12 months prior to the survey (Bendavid et al. 2011) are chosen for applying the Census method. The periods of the latest available successive censuses, WHO estimates (2011), and DHS surveys (Macro International 2012) are listed in Table 2. These periods are not identical. Nonetheless, noticing the limited improvements in the socioeconomic conditions in African countries during 1990-2010, comparing old-age mortality estimates in these periods may still be relevant.

Table 2. The periods of old-age mortality estimates

	Date of census[1]	Date of census[2]	WHO estimates	DHS periods
Burkina Faso	10/12/1996	09/12/2006	2000-2009	2007
Cote d'Ivoire	30/04/1975	01/05/1988	2000-2009	2005
Ethiopia	11/10/1997	29/05/2007	2000-2009	2008
Namibia	21/10/1991	27/08/2001	2000-2009	2006-2007
Nigeria	26/11/1991	21/03/2006	2000-2009	2008
Swaziland	11/03/1997	11/03/2007	2000-2009	2006-2007
Tanzania	28/08/1988	29/10/2002	2000-2009	2007-2008
Uganda	12/01/1991	12/09/2002	2000-2009	2006
Zambia	20/08/1990	25/10/2000	2000-2009	2007
Zimbabwe	18/08/1992	17/08/2002	2000-2009	2005-2006

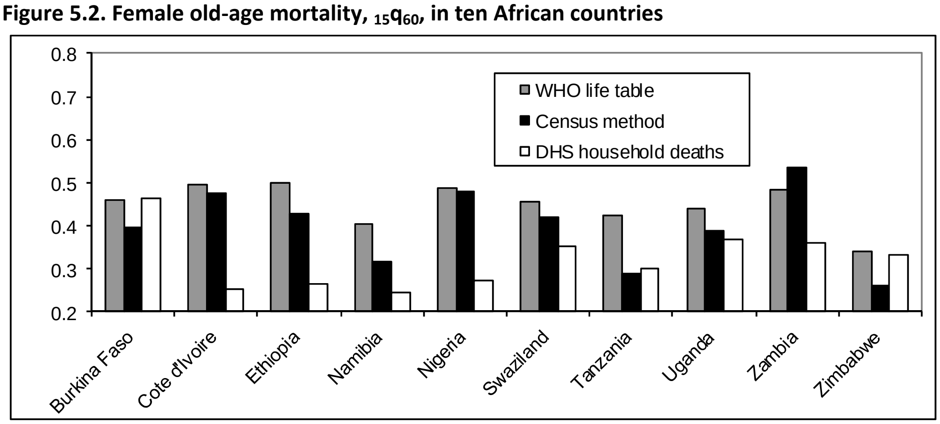
The comparisons are shown in Figure 5. It can be seen that, among these 10 African countries, old-age mortality obtained from the Census method are lower than those from WHO for 9 countries for males and females. The main reason of these systematic differences may be caused by the impact of HIV/AIDS, which may not directly affect old-age mortality, but significantly raises mortality at young and adult ages that are used to derive old-age mortality by WHO.

Figure 5.1. Male old-age mortality, $_{15}q_{60}$, in ten African countries



Source: computations by authors

Figure 5.2. Female old-age mortality, $_{15}q_{60}$, in ten African countries



Source: computations by authors

When comparing the estimates based on the Census method and those from DHS household death data, however, there are no systematic differences. It can be seen that among the 10 countries, old-age mortality obtained from the Census method are higher than that of using DHS data for 5 countries for males and for 7 countries for females. Besides that the DHS tends to underestimates mortality due to sample the deaths only in households with a woman in reproductive ages, the main reason for the random difference may remain in the limited sample size with regarding to old-age population and death. Overall, mortality rates based on unadjusted DHS household data are more often inconsistent with the Gompertz law (see Preston et al. 2001), which indicates that the growth rate of mortality increasing with age should be approximately constant.

Let the death rate at ages [x,x+5) be ${}_5m_x$ and the probability of dying between age x and x+5 be ${}_5q_x$. The values of ${}_5m_x$ can be computed by the estimated ${}_5q_x$ as ${}_5q_x/(5-2.5\cdot{}_5q_x)$ (see Preston et al. 2001). Subsequently, the growth rate of mortality increasing with age in the Gompertz law, namely g(x), can be computed as $g(x)=[Log\frac{{}_5m_{x+5}}{{}_5m_x}]$. According to the Gompertz law, the ratio of g(60)/g(65) should be close to 1. But the study using DHS data yields the values of this ratio that differ randomly and significantly from 1, as are shown in Table 3, which usually reflect insufficient sample size.

Table 3. Values of g(60)/g(65)

	Males	Females
Burkina Faso	0.08	0.17
Code d'Ivoire	0.70	-0.80
Ethiopia	1.27	1.78
Namibia	4.41	0.67
Nigeria	0.44	0.54
Swaziland	-5.32	-0.67
Tanzania	4.56	-4.69
Uganda	-66.07	0.73
Zambia	-0.05	-1.07
Zimbabwe	16.51	-0.12

Besides DHS, data from another survey, INDEPTH Network (2002), were also used to estimate old-age mortality. But the differences between the results from different sites in the same country are often too large, raising questions about whether they are representative at national level. Taking results from Tanzania females, as an example, the Hai site for 1992-1999 period showed a value of old-age mortality as ${}_{15}q_{60}$ = 0.24, but the Ifakara site for 1997-1999 indicated the value as 0.38 despite similar ${}_{45}q_{15}$ for both sites (${}_{45}q_{15}$ =0.29).

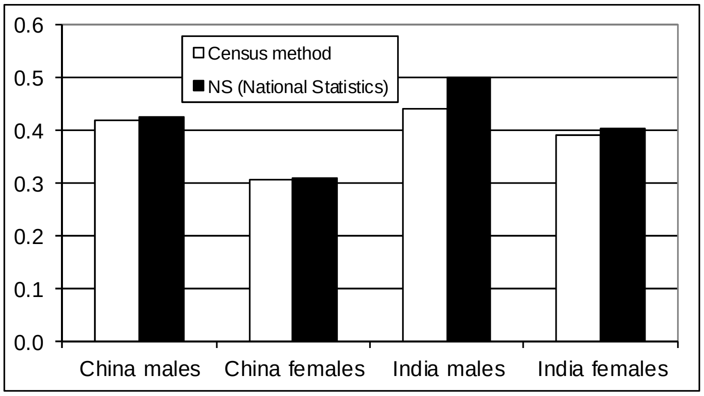
It should be mentioned clearly that the estimates of the Census method are not error free, and the main source of the errors is the difference in completeness of the two successive censuses. For example, old-age mortality seems to be underestimated for the females of Zimbabwe, compared to males of the same country and females from other countries. But, it is also clear that the Census method provide relevant information on old-age mortality for African countries.

Applications to Asian countries

In this section, we apply the Census method to the two most populous countries—China and India. As seen through these two examples, even in instances when old-age mortality estimates can be derived from direct data sources, the indirect Census method is helpful to improve our knowledge about old-age mortality.

In China, reliable death registrations are not yet widely available, and the numbers of death are collected from answers to retrospective questions in censuses. Although age heaping is not an issue due to the systematic use of the zodiac sign calendar, underreporting of death and population has been the general concern. Using the populations of the 1990 (July 1) and 2000 (November 1) censuses, old-age mortality in 1990-2000 can be estimated and shown in Figures 6. On the other hand, old-age mortality in 1989-1999 can also be estimated as averages between that of 1989 and 1999 life tables, computed directly from the census population and death (source: The National Bureau of Statistics of China), as are shown also in Figure 6 marked by NS (National Statistics). It can be seen that the two estimates are close for both males and females. These results indicate that there were no serious underreporting of death and population at old ages in the censuses; and reason may be that underreporting is more likely to appear among mobile populations, which are rare at old ages.

Figure 6. Old-age mortality, $_{15}q_{60}$, by sex based on 1990-2000 China and 1991-2001 India censuses



Source: computations by authors

For India, reliable national death registration is also not yet available, and the nationally representative statistical information about mortality is collected through the Sample Registration System (SRS), which usually only covers a small proportion of the whole population (Mahapatra 2010). Using the life tables obtained from the Sample Registration System (Source: Office of the Registrar) for 1991 and 1997, old-age mortality for 1991-1997 can be computed as the corresponding averages, as are shown in Figure 6 marked by NS (National Statistics). Like other surveys, representativity and completeness issues are the most common concerns. Moreover, age-heaping has also been an issue in Indian censuses (Gerland 2013). Using the Census method on the 1991 (March 1) and 2001 (March 1) censuses, the effect of age heaping is adjusted, and results are also displayed in Figure 6. It can be seen that the estimates of the Census method are lower (respectively -4% for females and -12% for males). These differences between national estimates based on SRS and indirect estimated based on the Census method might be due to improvement in the completeness of census over time, or by representativity issue with the SRS. Still, it is encouraging that these differences remain overall moderate between the two methods using entirely independent data.

Summary

Based on the fact that migration is rare at old ages, this paper proposed a new indirect method, namely the Census method, to estimate old-age mortality using three models and census data on population. The first model is the well-known variable-r method that converts the census populations into the person-years of the underlying stationary population. The second model uses the common relationship between survival ratios from model life tables to provide a robust way to adjust the effects of age-reporting errors in censuses. The third model is the extended Gompertz model, which estimates the starting-age population in a way consistent with commonly observed mortality patterns. The Census method could provide a chance to estimate old-age mortality, in principle, for all the countries that conducted more than one population census.

It should be mentioned, however, that under the current model specification presented in this paper for the Census method, the assumption that the completeness between successive censuses is comparable remains necessary. If the earlier census' completeness is lower than that of the later, then the Census method will underestimate old-age mortality, and vice versa. To gain more solid knowledge on old-age mortality for developing countries, it is clear that comparisons with results obtained from other methods and data are necessary.

Supplementary Appendices

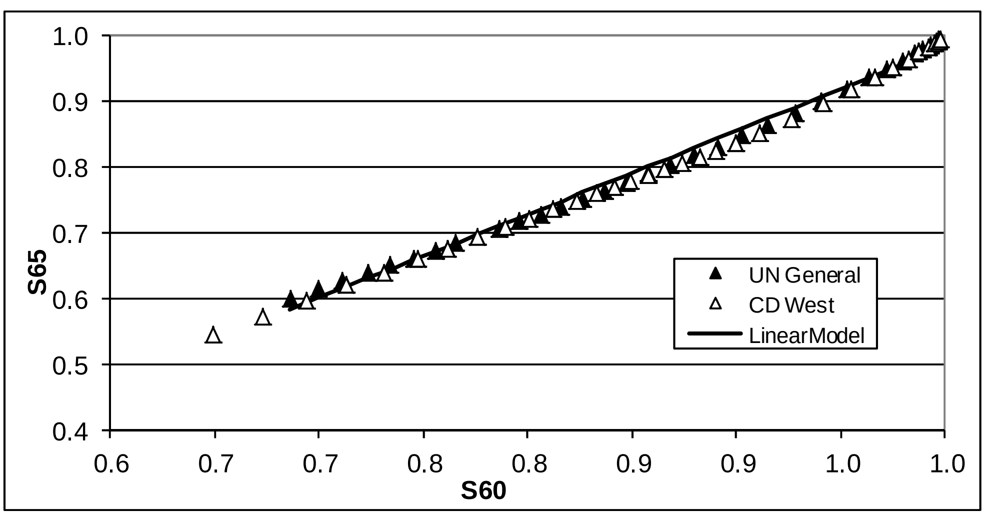
S1. The basis for adjusting errors in L_x

In a stationary population, denote the ratios of surviving from ages 60-64 to 65-69 by S_{60}, and from ages 65-69 to 70-74 by S_{65}:

$$S_{60} = \frac{L_{65}}{L_{60}}, \quad S_{65} = \frac{L_{70}}{L_{65}}. \tag{A.1}$$

The relationships between S_{60} and S_{65} can be explored in model life tables, which describes the common age patterns of mortality. Using the United Nations General (UN General) model life tables (United Nations 1982) that are based on mortality data from developing countries, and the Coale-Demeny West (CD_W) model life tables (Coale, Demeny and Vaughan 1983) that are established on the data of developed countries, the relationships for females are shown in Supplementary Figure S1, in which the levels of life expectancy are from 20 to 100 year using the extended model life tables computed by the United Nations (2010) and Li and Gerland (2011).

Figure S1. Relationship between female survival ratios from age 60-64 to 65-69 (S_{60}), and 65-69 to 70-74 (S_{65}) for selected model life table families



Source: computations by authors based on United Nations (2010)

Applying a linear model,

$$S_{65} = a + b \cdot S_{60}, \qquad \text{(A.2)}$$

to describe the relationships from the UN general family in Figure 1, results are:

$$a = -0.29, \quad b = 1.27, \quad R^2 = 0.997, \qquad \text{(A.3)}$$

where R^2 is the explanation ratio, which indicates that, for the female UN general model life tables at the levels of life expectancy from 20 to 100 years, 99.7% of variance in S_{65} is explained by model (A.2)-(A.3). Using (A.2)-(A.3) and S_{60} to predict the S_{65} of other model life table families at the levels of life expectation from 20 to 100 years, the values of R^2 are listed in Table 1.

Table S1. The explanation ratios (R^2) of applying equation (A.2) with (A3) parameters (estimated for females UN general model) to the CD and UN model life tables

	CD East	CD North	CD South	CD West	UN Chilean	UN Far East Asian	UN General	UN Latin American	UN South Asian
Males	0.989	0.993	0.985	0.993	0.994	0.988	0.996	0.996	0.996
Females	0.981	0.987	0.971	0.993	0.995	0.986	0.997	0.994	0.993

Model (A.2)-(A.3) is established on female data of the UN general family that is based on mortality data of developing countries, and it successfully described the relationships between S_{60} and S_{65} for males and for other model life table families, including the CD families that are based on mortality data from developed countries. Thus, (A.2)-(A.3) is the simplest description of a relationship between S_{60} and S_{65}, which is robust to the levels and patterns of mortality, and robust to the mortality of developing countries or developed countries. Since this relationship is robust, it does not matter which model, family or sex is used to estimate the values of parameters as are shown in (A.3); and (A.2)-(A.3) could be used as a basis to adjust the errors in the L_x. For further reference, a full set of values for all model life tables is also provided in Table S2.

Table S2. Linear regression coefficients for equations (A.2) for CD and UN model life tables

	CD East	CD North	CD South	CD West	UN Chilean	UN Far East Asian	UN General	UN Latin	UN South Asian
Females									
Intercept (a)	-0.38	-0.39	-0.44	-0.30	-0.28	-0.22	-0.29	-0.31	-0.31
Slope (b)	1.37	1.38	1.42	1.28	1.26	1.20	1.27	1.30	1.29
R^2	0.998	0.999	0.998	0.997	0.997	0.997	0.997	0.997	0.997
Males									
Intercept (a)	-0.37	-0.37	-0.40	-0.27	-0.26	-0.22	-0.28	-0.31	-0.30
Slope (b)	1.35	1.36	1.38	1.25	1.25	1.20	1.27	1.30	1.28
R^2	0.998	0.999	0.999	0.997	0.998	0.997	0.998	0.998	0.998

In describing the age patterns of mortality change, model life tables that cover all the age range are proved necessary. But why the relationships between S_{60} and S_{65} can be so simple? The main reason is that the range of ages related to S_{60} and S_{65} is not large, and in this range the changes of mortality are simple (as is described by the Gompertz law), and are highly correlated (United Nations 1982).

S2. The adjustment dealing with age heaping errors in L_x

Age heaping at 60 and 70 years of age is common in censuses of developing countries. Although using 5-year age group eliminates errors caused by those older than 70, for example, among those claiming their age as 70, errors caused by people aged 65-69 reporting their age as 70 still exist. When such age heaping appears in successive censuses, populations in age group 65-69 would be fewer than those in 70-74 in two censuses, and the L_{65} would be likely fewer than L_{70}, which is unacceptable.

The effect of age heaping on 70 can be described as an amount of Δ persons aged 65-69 claiming their age as 70 years, and hence counted in L_{70}. Denoting the heaping ratio at age 70 as $\frac{\Delta}{L_{70}}$, the number of heaped person at age 70 is $L_{70}(\frac{\Delta}{L_{70}})$. Assuming that the heaping ratio at age 60 equals that at age 70, then the number of heaped person at age 60 is $L_{60}(\frac{\Delta}{L_{70}})$. Thus, the age-heaping adjustment is written as

$$\begin{aligned} \hat{L}_{60} &= L_{60} - \frac{L_{60}}{L_{70}}\Delta, \\ \hat{L}_{65} &= L_{65} + \Delta, \\ \hat{L}_{70} &= L_{70} - \Delta. \end{aligned} \tag{A.4}$$

How to estimate Δ? This paper uses the simplest description of the robust relationship, (A.2)-(A.3), as the target for the adjustment to reach. In other words, Δ is chosen to fit model (A.2)-(A.3) as below:

$$\frac{L_{70} - \Delta}{L_{65} + \Delta} = a + b\frac{L_{65} + \Delta}{L_{60} - \frac{L_{60}}{L_{70}}\Delta}\ . \tag{A.5}$$

Equation (A.5) leads to

$$\begin{aligned} \Delta &= \frac{-B + \sqrt{B^2 - 4AC}}{2A}, \\ A &= b - a\frac{L_{60}}{L_{70}} - \frac{L_{60}}{L_{70}}, \\ B &= a(L_{60} - \frac{L_{60}}{L_{70}}L_{65}) + 2bL_{65} + L_{60} + \frac{L_{60}}{L_{70}}L_{70}, \\ C &= L_{65}(aL_{60} + bL_{65}) - L_{60}L_{70}. \end{aligned} \tag{A.6}$$

In (A.6), A and B are always positive; and C is negative when point (S_{60}, S_{65}) is above the model line in Figure S1, which implies age heaping on 60 and 70 years. Thus, when point (S_{60}, S_{65}) is above the model line, there is $\Delta > 0$, and the effect of age heaping on L_x is adjusted.

But when point (S_{60} , S_{65}) deviates from the model line in Figure S1, it is not necessarily a result of age heaping, especially when the deviation is below the model line. In this paper, age heaping is identified only for abnormal situations $S_{60} < S_{65}$, and other deviations from the model line are adjusted by a minimal adjustment as below (see Supplementary Appendix S3).

S3. The minimal adjustment dealing with nonspecific errors in L_x

The errors in L_x may have multiple reasons, such as undercounting specific subgroups of population in one census, but not in another. In this paper, a minimal adjustment is proposed to deal with errors caused by nonspecific reasons. The first step of this minimal adjustment is to adjust survival ratios as $\hat{S}_{60}$ and $\hat{S}_{65}$, which fit (A.2)-(A.3). Because point ($\hat{S}_{60}$, $\hat{S}_{65}$) must be on the model line, the adjustment is made by choosing a $\hat{S}_{60}$ to minimize $[(\hat{S}_{60} - S_{60})^2 + (a + b\hat{S}_{60} - S_{65})^2]$. And the solution of the minimization is

$$\hat{S}_{60} = \frac{-ab + S_{60} + bS_{65}}{1 + b^2}, \quad \hat{S}_{65} = a + b\hat{S}_{60}. \tag{A.7}$$

On the plane as is shown in Figure S1, the adjustment minimizes the distance between point (S_{60} , S_{65}) and the model line. Thus, the segment connecting points (S_{60} , S_{65}) and ($\hat{S}_{60}$, $\hat{S}_{65}$) must be orthogonal to the model line. When the point (S_{60} , S_{65}) is below the model line, for example, the orthogonal property indicates that the adjustment reduces S_{60} , and increases S_{65} . Because the adjustments on S_{60} and S_{65} cancel each other, their effect on old-age mortality, ${}_{15}q_{60}$, will be small.

The second step is to adjust the L_x to fit the adjusted survival ratios. Let the adjusted L_x be $\hat{L}_x$, and notice that $\hat{L}_x$ must satisfy $\hat{S}_x$, the adjustment is made by choosing a $\hat{L}_{60}$ to minimize

$$\begin{aligned}&[(\hat{L}_{60} - L_{60})^2 + (\hat{L}_{65} - L_{65})^2 + (\hat{L}_{70} - L_{70})^2] = \\ &[(\hat{L}_{60} - L_{60})^2 + (\hat{S}_{60}\hat{L}_{60} - L_{65})^2 + (\hat{S}_{60}\hat{S}_{65}\hat{L}_{60} - L_{70})^2].\end{aligned} \tag{A.8}$$

The solution of this minimization is

$$\hat{L}_{60} = \frac{L_{60} + \hat{S}_{60}L_{65} + \hat{S}_{60}\hat{S}_{65}L_{70}}{1 + \hat{S}_{60}^2 + \hat{S}_{60}^2\hat{S}_{65}^2}, \quad \hat{L}_{65} = \hat{S}_{60}\hat{L}_{60}, \quad \hat{L}_{70} = \hat{S}_{65}\hat{L}_{65}. \tag{A.9}$$

Since the reason for doing this minimal adjustment is not as strong as that of adjusting age heaping, we recommend to use this minimal adjustment in a weighted way:

$$\hat{L}_{60} = w\frac{L_{60} + \hat{S}_{60}L_{65} + \hat{S}_{60}\hat{S}_{65}L_{70}}{1 + \hat{S}_{60}^2 + \hat{S}_{60}^2\hat{S}_{65}^2} + (1-w)L_{60},$$

$$\hat{L}_{65} = w\hat{S}_{60}\frac{L_{60} + \hat{S}_{60}L_{65} + \hat{S}_{60}\hat{S}_{65}L_{70}}{1 + \hat{S}_{60}^2 + \hat{S}_{60}^2\hat{S}_{65}^2} + (1-w)L_{65}, \qquad \text{(A.10)}$$

$$\hat{L}_{70} = w\hat{S}_{65}\hat{S}_{60}\frac{L_{60} + \hat{S}_{60}L_{65} + \hat{S}_{60}\hat{S}_{65}L_{70}}{1 + \hat{S}_{60}^2 + \hat{S}_{60}^2\hat{S}_{65}^2} + (1-w)L_{70},$$

where $0 \le w \le 1$ is the weight, and is used as 0.5 in this paper.

References:

Bendavid, E., B. Seligman, and J. Kubo. 2011. "Comparative Analysis of Old-Age Mortality Estimations in Africa." *PLoS ONE* 6(10):e26607.

Bennett, N.G.and S. Horiuchi. 1981. "Estimating the Completeness of Death Registration in a Closed Population." *Population Index* 47(2):207-221.

Brass, W. 1964. "Uses of census or survey data for the estimation of vital rates." Presented at UN. ECA African Seminar on Vital Statistics, Dec.14 - 19, Addis Ababa, Ethiopia.

Brass, W.and A.J. Coale. 1968. "Methods of Analysis and Estimation." Pp. 88-139 in *The Demography of Tropical Africa*, edited by W. Brass, A.J. Coale, P. demeny, D.F. Heisel, F. Lorimer, A. Romaniuk, and E. Van de Walle. Princeton, NJ: Princeton University. Office of population research. Princeton University Press.

Coale, A.J.and P.G. Demeny. 1966. *Regional model life tables and stable populations*. Princeton, N.J.,: Princeton University Press.

Coale, A.J., P.G. Demeny, and B. Vaughan. 1983. *Regional model life tables and stable populations*. New York: Academic Press.

Ewbank, D.C. 1981. *Age misreporting and age-selective underenumeration : sources, patterns, and consequences for demographic analysis*: National Academy Press.

Gerland, P. 2013. "UN Population Division's Methodology in Preparing Base Population for Projections: case study for India." Pp. 22 in *JYP Comparative Asia Research Centre, NUS Global Asia Institute Conference on "40 per cent of the world: population change, human capital and development in China, India and Indonesia"*. Singapore: National University of Singapore.

Hill, K. 1987. "Estimating census and death registration completeness." *Asian Pac Popul Forum* 1(3):8-13, 23-14.

INDEPTH Network. 2002. *Population and health in developing countries. Volume 1. Population, health, and survival at INDEPTH sites*. Ottowa, Canada: International Development Research Centre.

Li, N.and P. Gerland. 2011. "Modifying the Lee-Carter Method to Project Mortality Changes up to 2100." in *Annual Meeting of the Population Association of America*. Washington, DC: Population Association of America.

Macro International. 2012. "MEASURE DHS. Demographic and Health Surveys." Calverton (Maryland): Available: http://www.measuredhs.com/.

Mahapatra, P. 2010. "An Overview of the Sample Registration System in India." in *Prince Mahidol Award Conference & Global Health Information Forum*.

Murray, C.J., J.K. Rajaratnam, J. Marcus, T. Laakso, and A.D. Lopez. 2010. "What can we conclude from death registration? Improved methods for evaluating completeness." *PLoS Med* 7(4):e1000262.

Murray, C.J.L., B.D. Ferguson, A.D. Lopez, M. Guillot, J.A. Salomon, and O. Ahmad. 2003. "Modified Logit Life Table System: Principles, Empirical Validation, and Application." *Population Studies* 57(2):165-182.

Preston, S.H.and A.J. Coale. 1982. "Age Structure, Growth, Attrition, and Accession: A New Synthesis." *Population Index* 48(2):217-259.

Preston, S.H., I.T. Elo, and Q. Stewart. 1999. "Effects of age misreporting on mortality estimates at older ages." *Population Studies-a Journal of Demography* 53(2):165-177.

Preston, S.H., P. Heuveline, and M. Guillot. 2001. *Demography : measuring and modeling population processes*. Malden, MA: Blackwell Publishers.

Rajaratnam, J.K., J.R. Marcus, A. Levin-Rector, A.N. Chalupka, H. Wang, L. Dwyer, M. Costa, A.D. Lopez, and C.J. Murray. 2010. "Worldwide mortality in men and women aged 15-59 years from 1970 to 2010: a systematic analysis." *Lancet* 375(9727):1704-1720.

Timæus, I.M. 2013a. "Indirect estimation from orphanhood in multiple inquiries." Pp. 293-308 in *Tools for Demographic Estimation*, edited by T.A. Moultrie, R.E. Dorrington, A.G. Hill, K. Hill, I.M. Timæus, and B. Zaba: International Union for the Scientific Study of Population.

—. 2013b. "Indirect estimation of adult mortality from orphanhood." Pp. 222-243 in *Tools for Demographic Estimation*, edited by T.A. Moultrie, R.E. Dorrington, A.G. Hill, K. Hill, I.M. Timæus, and B. Zaba: International Union for the Scientific Study of Population.

United Nations. 1982. *Model life tables for developing countries*. New York: United Nations.

—. 1983. *Manual X: indirect techniques for demographic estimation*. New York: United Nations.

—. 2002. *Methods for estimating adult mortality*. New York: United Nations.

—. 2010. "World Population Prospects: The 2010 Revision - WPP2010 Extended Model Life Tables." New York: Dept. of Economic and Social Affairs, Population Division.

—. 2012. *Population Ageing and Development 2012 (Wall Chart)*. New York, N.Y.: United Nations Publications.

—. 2013. "World Population Prospects: The 2012 Revision." New York: Dept. of Economic and Social Affairs, Population Division.

United Nations. Statistical Division. 1948-. "Demographic Yearbook - Annuaire Démographique." edited by D. United Nations Statistics Division. New York, NY: Department of Economic and Social Affairs, Statistical Office, United Nations.

Wang, H., L. Dwyer-Lindgren, K.T. Lofgren, J.K. Rajaratnam, J.R. Marcus, A. Levin-Rector, C.E. Levitz, A.D. Lopez, and C.J. Murray. 2012. "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.

WHO. 2011. "Life tables." edited by G.H.O. (GHO). Geneva: World Health Organization.

Wilmoth, J., S. Zureick, V. Canudas-Romo, M. Inoue, and C. Sawyer. 2011. "A flexible two-dimensional mortality model for use in indirect estimation." *Population Studies* 66(1):1-28.

Using census data to estimate old-age mortality for developing countries

Nan Li and Patrick Gerland
(gerland@un.org)

IUSSP 2013 – Busan, Rep. of Korea (27 Aug. 2013)
Session 46: Indirect methods of mortality and fertility estimation: new techniques for new realities

United Nations Population Division
Estimates and Projection Section

Outline

- Background
- Model and estimation strategy
- Application

Background

Background

- UN *World Population Prospects: 2012 Revision* (see meta-information)
- Out of 156 countries with 90,000 inhabitants or more in 2013, used **vital registration data on mortality for only 70 of them (45%)**
- For the rest, relied on censuses and surveys

Background

- United Nations (2013). *World Mortality Report 2013: Data Inventory*
- *Out of* ***163 countries that used censuses or surveys to collect mortality data between 1948-2013***
  - *Greater knowledge and data availability on child mortality*
  - *More limited and partial info on adult mortality*

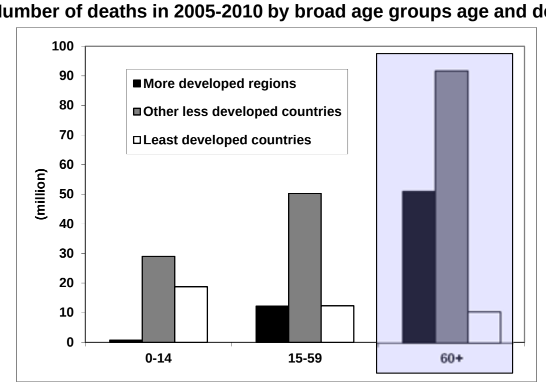
	Child mortality			Adult mortality				
	CEB/CS	Maternity histories	HH deaths	Maternal orphan.	Paternal orphan.	Siblings survival	Widowhood	n
Surveys	707	537	129	68	54	178	24	879
Census	426	0	167	115	84	2	5	440
Surveys	80%	61%	15%	8%	6%	20%	3%	100%
Census	97%	0%	38%	26%	19%	0%	1%	100%

Background

- With improvements in child and adult survival in last decades, **old-age mortality** has become the elephant in the room

Figure 1. Number of deaths in 2005-2010 by broad age groups age and development group

55% of deaths worldwide in 2005-10 at age 60+
(153 out of 277 million deaths)



More than half of all deaths occur at age 60+ in nearly all developing regions
(except SSA with only one fourth)

Research questions

- To estimate old-age mortality in developing countries lacking VR, the prevailing practice relies either on:
  - 5q0 and model life tables
  - 5q0 and 45q15 and some matching model pattern with a relational logit transformation or PCA on the first 1-3 components
- Our proposal is to infer old-age mortality based on data on old-age population

Old-age mortality data from censuses and surveys

- Indirect estimates based on survival of relatives are limited in scope, and often do not include older people and their reliability often problematic
- Direct estimates based on HH deaths from censuses are limited, and from surveys even more rare and less reliable due to sample size limitations and sample selection (Bendavid, Seligman and Kubo 2011)

Proposal

- Estimate indirectly old-age mortality using data from population censuses
  (over the last 60 years, 87% of countries have ≥ 3 censuses, 77% have ≥ 4 and about half of the countries have ≥ 5)
  - Addressing issues related to errors of age reporting
  - Assuming difference in completeness between successive censuses is negligible
- No standard measure of old-age mortality
  - Use the probability of dying between ages 60 and 75 ($_{15}q_{60}$) based on data availability for open age group

Model and estimation strategy

Census method to estimate $_{15}q_{60}$

$$_{15}q_{60} = 1 - \frac{l_{75}}{l_{60}}$$

l_x is the number of survivors at age x

- Approach in 3 steps:
  - Variable-r method
  - Adjustment of errors on age
  - Extended Gompertz model

1. Variable-r method (Preston et al. 2001)

- Use the population from two censuses, and compute for the intercensal period the person-years lived in 5-year age group *[x,x+5)* in the underlying stationary population

$$L_x = N(x)\exp[s(x)], x = 60,65,70.$$

with intercensal mid-population in age group *[x,x+5)*

$$N(x) = \sqrt{p(x,t_1)\,p(x,t_2)}, \quad x = 60,65,70$$

p(x, t) = observed population in age group [x,x+5) in a census at time t

and s(x) accumulated growth rates

$$s(60) = 2.5r(60),$$
$$s(65) = 5r(60) + 2.5r(65),$$
$$s(70) = 5[r(60) + r(65)] + 2.5r(70)$$

2. Adjustment of errors on age

- variable-r assumptions:
  - Migrations are controlled for (or don't matter)
  - Errors in enumerating population are constant with age
  - Errors are identical in the two successive censuses
- But in developing countries the errors often occur unevenly across age: e.g., age heaping and age exaggeration

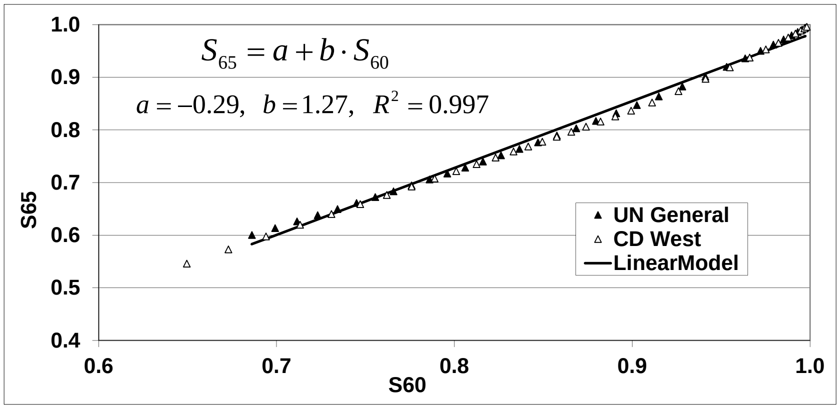
2. Adjustment of errors on age

- Adjust age errors for stationary populations a common relationship between the survival ratios $S_{60}=L_{65}/L_{60}$ and $S_{65}=L_{70}/L_{65}$ and among model life tables, which is

$$S_{65} = -0.29 + 1.27 * S_{60}$$

  based on the UN general family for females.
- This relationship is used as model line, and serve as basis to adjust age errors and compute adjusted L_x (see paper for details) $\hat{L}_{60}, \hat{L}_{65}, \hat{L}_{70}$

Relationship between female survival ratios from age 60-64 to 65-69 (S_{60}), and 65-69 to 70-74 (S_{65}) for selected model life table families



Explanation ratios (R^2) of applying the UN general model to other model life tables

	CD East	CD North	CD South	CD West	UN Chilean	UN Far East Asian	UN General	UN Latin American	UN South Asian
Males	0.989	0.993	0.985	0.993	0.994	0.988	0.996	0.996	0.996
Females	0.981	0.987	0.971	0.993	0.995	0.986	0.997	0.994	0.993

2. Adjustment of errors on age

When observed survival-ratio point, (S_{60},S_{65}), is:

- Above the model line, or when the survival ratio is abnormally rising with age, the difference between the survival-ratio point and the model line is caused mainly by **age heaping**
- Below the model line, the difference between the survival-ratio point and the model line is caused by **nonspecific errors**

3. Extended Gompertz model

- variable-r method produces only L_x, but l_x still unknown and must be estimated in a way consistent with observed mortality patterns
- Gompertz law: at older ages, force of mortality $\mu(x) = -\frac{dl_x}{l_x dx}$ rises roughly exponentially

$$\mu(x) = \mu(60)\exp[g(x-60)], \quad g > 0$$

μ(60) is the force of mortality at age 60 and *g* the rate of mortality increase

- Extended Gompertz model includes an additional parameter, l_{60}

$$l_x = l_{60}\exp[-\int_{60}^{x}\mu(y)dy] = l_{60}\exp\{-\frac{\mu(60)}{g}[\exp[g(x-60)]-1]\}$$

3. Extended Gompertz model

- Three unknown parameters namely l_{60}, *μ(60)*, and *g*; and three equations to solve these parameters according to L_x

$$\hat{L}_{60} = \int_{60}^{65} l_x dx = l_{60}\int_{60}^{65}\exp\{-\frac{\mu(60)}{g}[\exp[g(x-60)]-1]\}dx,$$

$$\hat{L}_{65} = l_{60}\int_{65}^{70}\exp\{-\frac{\mu(60)}{g}[\exp[g(x-60)]-1]\}dx,$$

$$\hat{L}_{70} = l_{60}\int_{70}^{75}\exp\{-\frac{\mu(60)}{g}[\exp[g(x-60)]-1]\}dx.$$

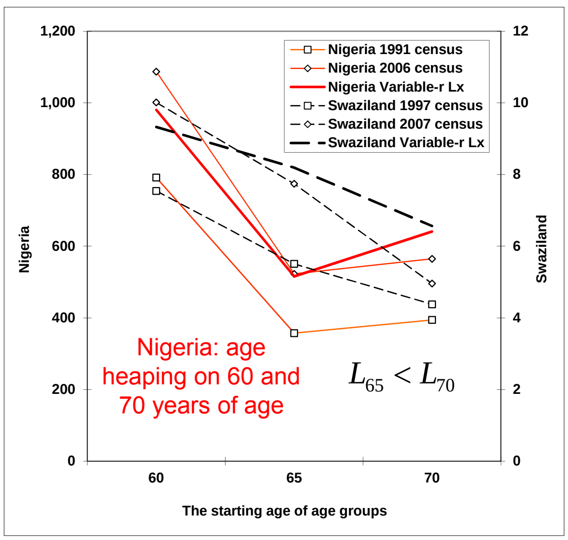
- Solve using non-linear optimization to get estimate of old-age mortality as:

$${}_{15}q_{60} = 1 - \frac{l_{75}}{l_{60}} = 1 - \exp\{-\frac{\mu(60)}{g}[\exp[g(75-60)]-1]\}$$

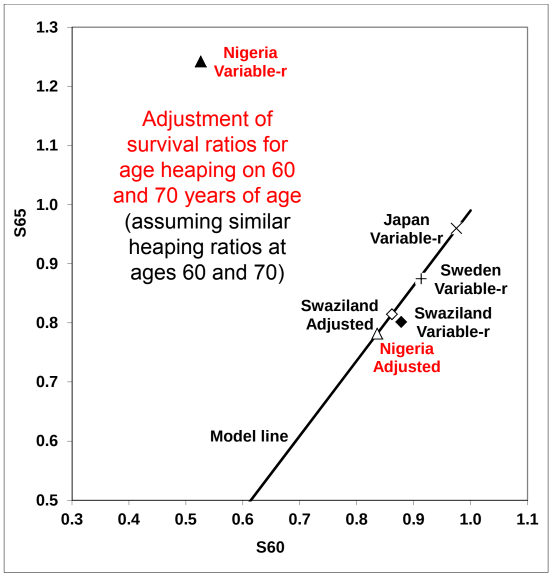
Application

Female census population (UN DYB): Japan (2000, 2005), Sweden (1950,1960), Nigeria (1991,2006, Swaziland (1997,2007)

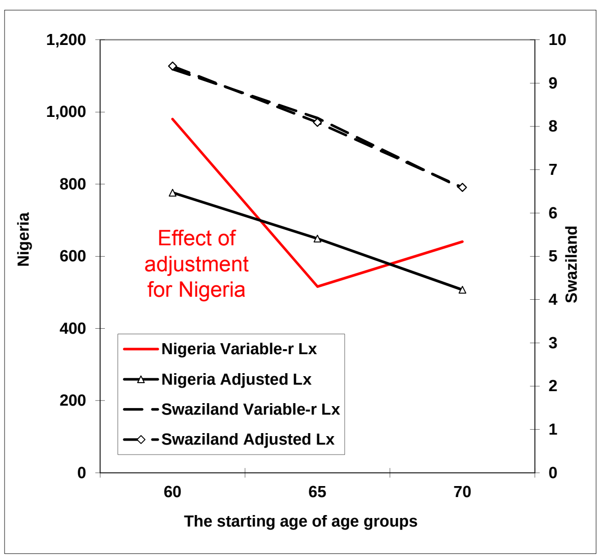
Enumerated female population in 5-year age group for age 60 and over in 1991-2006 Nigeria and 1997-2007 Swaziland censuses (in thousands)



Variable-r and adjusted survival ratios of females for selected countries



Variable-r (un)adjusted Lx in 5-year age group for female age 60 and over in 1991-2006 Nigeria and 1997-2007 Swaziland censuses (in thousands)



Results: estimates of $_{15}q_{60}$

- Estimates of old-age mortality based on applying the extended Gompertz model to the adjusted L_x

	Period	Males	Females
Japan	2000-2005	0.214	0.099
Sweden	1960-1965	0.370	0.292
Nigeria	1991-2006	0.356	0.479
Swaziland	1997-2007	0.624	0.419

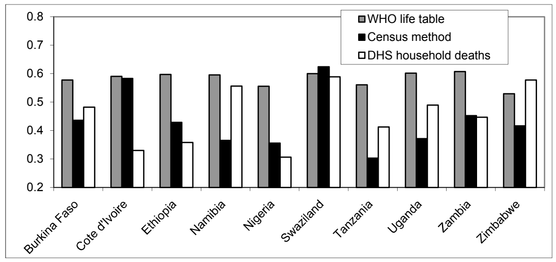
Difference in completeness between the two censuses (e.g., greater improved completeness for males than for females) ?

Application to 10 DHS African countries (with HH deaths in last 12 months)

Periods of old-age mortality estimates

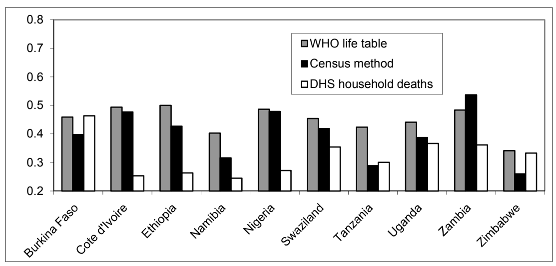
	Date of census[1]	Date of census[2]	WHO estimates	DHS periods
Burkina Faso	10/12/1996	09/12/2006	2000-2009	2007
Cote d'Ivoire	30/04/1975	01/05/1988	2000-2009	2005
Ethiopia	11/10/1997	29/05/2007	2000-2009	2008
Namibia	21/10/1991	27/08/2001	2000-2009	2006-2007
Nigeria	26/11/1991	21/03/2006	2000-2009	2008
Swaziland	11/03/1997	11/03/2007	2000-2009	2006-2007
Tanzania	28/08/1988	29/10/2002	2000-2009	2007-2008
Uganda	12/01/1991	12/09/2002	2000-2009	2006
Zambia	20/08/1990	25/10/2000	2000-2009	2007
Zimbabwe	18/08/1992	17/08/2002	2000-2009	2005-2006

Male old-age mortality, $_{15}q_{60}$, in ten African countries



Census method lower than WHO for 9 countries

Female old-age mortality, $_{15}q_{60}$, in ten African countries



No systematic differences between census method and DHS HH deaths

- Growth rate of mortality increasing with age in the Gompertz law, namely g(x), can be computed as…

$$g(x) = [Log \frac{_5m_{x+5}}{_5m_x}]$$

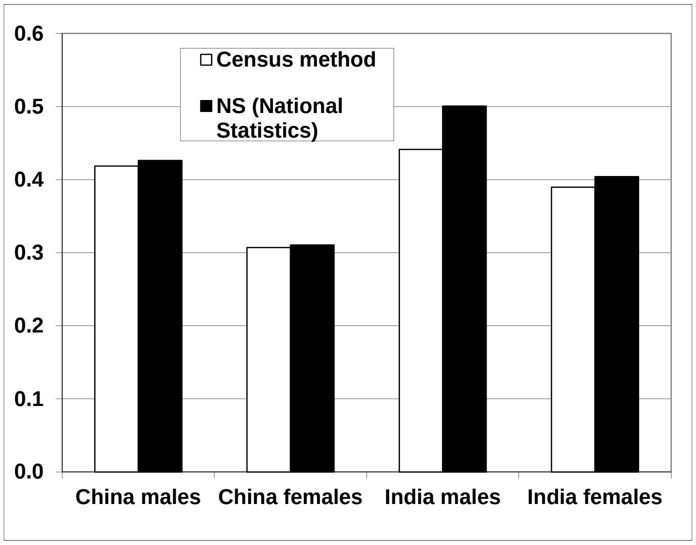
- According to the Gompertz law, the ratio of g(60)/g(65) should be close to 1

Ratios based on DHS HH deaths

	Males	Females
Burkina Faso	0.08	0.17
Code d'Ivoire	0.70	-0.80
Ethiopia	1.27	1.78
Namibia	4.41	0.67
Nigeria	0.44	0.54
Swaziland	-5.32	-0.67
Tanzania	4.56	-4.69
Uganda	-66.07	0.73
Zambia	-0.05	-1.07
Zimbabwe	16.51	-0.12

Ratios differ randomly and significantly from 1

Old-age mortality, $_{15}q_{60}$, by sex based on 1990-2000 China and 1991-2001 India censuses



Similar in China

For India, Census method lower (4% for females, -12% for males)

Discussion

- Assumption of similar completeness of two successive censuses still applies for this method.
- Under the assumption that census coverage improves over time, any improved completeness for the most recent census would lead to underestimation of old age mortality – unless the respective censuses are adjusted for differential completeness using PES or DDM.
- To improve robustness of intercensal estimates, test/use alternative intercensal periods using longer periods and multiple censuses.
- Evaluation of final old-age mortality results by sex can also reveal further insights about the plausibility of the underlying assumption about differential census coverage between censuses (e.g., male mortality substantially much lower than female old-age mortality could be one of the symptoms).