THE MODIFIED ORPHANHOOD METHOD: POTENTIALS AND LIMITATIONS

Marc Luy and Angela Wiedemann

Wittgenstein Centre (IIASA, VID/ÖAW, WU),
Vienna Institute of Demography of the Austrian Academy of Sciences

- Paper to be presented at the XXVII IUSSP International Population Conference in Busan, Korea -

Abstract

The work of demographers in developing countries is limited by the fact that demographic data are either nonexistent or of too bad quality to be usable. Therefore, indirect estimation techniques based on survey data are used to estimate levels and trends of mortality and life expectancy. But also mortality researchers in developed countries are often faced with the problem of nonexistent data when they are interested in specific phenomena of mortality differentials. In order to improve the availability of information on specific mortality differentials in cases where no official data are available, Luy (2009, 2012) suggested a modified version of the orphanhood method (MOM) for indirect estimation of adult mortality from survey information on maternal and paternal survival to permit its application to populations of developed countries. The aim of this paper is to demonstrate the formal demographic relationships behind the MOM and to illustrate the method’s potentials and limitations by summarizing the findings of different empirical applications. The latter will provide additional insights not only for the general usefulness of indirect estimation techniques in developed countries but also for the application of the traditional variants of the orphanhood method in developing countries.

Introduction

The work of demographers in developing countries is limited by the fact that demographic data are either nonexistent or of too bad quality to be usable. Therefore, indirect estimation techniques based on survey data are used to estimate demographic parameters, such as levels and trends of mortality and life expectancy. But also mortality researchers in developed countries are often faced with the problem of nonexistent data when they are interested in specific phenomena of mortality differentials. In view of these shortcomings and the continuous rise of surveys including the information for applying indi-
rect estimation techniques, the rare use of these methods outside the developing world is surprising. In order to improve the availability of information on specific mortality differentials in cases where no official data are available, Luy (2009, 2012) suggested a modified version of the orphanhood method for indirect estimation of adult mortality from survey information on maternal and paternal survival to permit its application to populations of developed countries.

In the meanwhile, the ‘modified orphanhood method’ (MOM) has been applied to estimate differences in life expectancy by education and occupation in Italy (Luy et al. 2011), by education in Germany (Luy et al. 2010) and between migrants and non-migrants in Germany (Wiedemann 2012). These empirical applications illustrate the practical applicability of this method as well as the general potential of indirect estimation techniques in developed countries. However, some results of these studies revealed also its limitations when case numbers are small or the survey data is of bad quality.

The aim of this paper is

1. to demonstrate the formal demographic relationships behind the MOM, and
2. to illustrate the method’s potentials and limitations by summarizing the findings of different empirical applications.

The latter will provide additional insights not only for the general usefulness of indirect estimation techniques in developed countries but also for the application of the traditional variants of the orphanhood method in developing countries.

**The modified orphanhood method (MOM)**

The orphanhood method is a demographic technique to estimate adult mortality from survey information about parents’ survival. Its basic idea is that the age of respondents represents the survival time of the mother (or father). The proportion of survey respondents of a given age group whose mother (or father) is still alive approximates a survivorship ratio from an average age at childbearing to that age plus the age of the respective respondents (see e.g. Preston et al. 2001). The traditional variants of the orphanhood method model this relation using different theoretical patterns of fertility, mortality and age composition in order to convert the share of those with a surviving parent into a life table survivorship probability, controlling for the actual pattern of childbearing (Brass and Hill 1973; Hill and Trussell 1977; Zlotnik and Hill 1981; Hill et al. 1983; Chackiel and Orellana 1985; Timæus 1986, 1991a, 1991b, 1992; Timæus and Nunn 1997).

In less developed countries, the use of theoretical population models for this conversion is necessary because there are no data on basic fertility and mortality patterns. However, for most populations of
developed countries, detailed age-specific fertility and mortality rates are available for both periods and cohorts. The MOM addresses this issue and converts the proportion of respondents of a given age group with mother/father still alive into period survivorship probabilities from age 30 on the exclusive basis of empirical data. A description of the logic of these transformation and the formal demographic details can be found in the appendix. Using Brass’ two-parameter relational logit model life table system (Brass 1971, 1975) with the official life table for the determined reference period as standard, the survivorship probabilities derived with the MOM can be transferred into complete life tables from age 30 and thus into estimates for life expectancy.

Data used in the empirical applications

In the studies mentioned above the MOM was applied to the cross-sectional data of the Italian Multi-purpose Surveys (MPS) of the years 1998 (ages 20-64, n= 38,227) and 2003 (ages 20-64, n=32,162), the German Integration Survey (INTSY) of the year 2000 (ages 18-29, N= 2,465), the German sample of the Gender and Generations Survey (GGS) of the year 2005 (ages 20-64, n=7,787), and the German sample of the Survey of Health, Aging and Retirement in Europe (SHARE) of the year 2004 (ages 50-64, n=1,573).

Results

MOM estimates for overall life expectancy in Italy with MPS data

Figure 1 shows the estimates for life expectancy at age 35 for Italian women and men from the late 1970s to the mid-1990s, based on data of the Italian MPS from 1998 (green dots) and 2003 (red dots). The single estimates fluctuate (stronger for men than for women) and therefore it is difficult to use them separately. However, it is possible to combine them to estimate mortality or life expectancy for specific periods (e.g. 1980-84, 1985-89 etc.) or to estimate functions for the time trend as illustrated in Fig. 1 with linear trend estimations. The graph reveals the applicability of the method since the time trends in life expectancy of the total Italian population (based on data from the Human Mortality Database HMD) are reflected very well by the MOM estimates. However, the life expectancy estimates derived with the MOM lie above the HMD values. This is due to the fact that the orphanhood method is by definition based exclusively on the survival experiences of parous (and in most cases also married) women and men whose lower mortality is known from many studies. This has only minor effects when the method is used to study mortality differentials since all population subgroups are similarly affected by this effect. However, this might be an important issue in developing countries where the orphanhood method is used to estimate overall mortality and life expectancy.
MOM estimates for overall life expectancy in Germany with German GGS data

The results shown in Figure 2 demonstrate that the applicability of the MOM is highly dependent on the quality of the survey data. Whereas the MOM estimates for female life expectancy in Germany compared to official population statistics provide a similar picture as the corresponding estimates for Italy in Figure 1 (note that the estimates for Italy are based on much higher case numbers than those for Germany), the MOM estimates for male life expectancy in Germany are misleading and indicate a strong decrease in life expectancy between the late 1980s and the mid-1990s. These obvious wrong results are, however, not due to the estimation procedure but to the quality of the German GGS data which is particularly problematic for young survey respondents (see Kreyenfeld et al. 2011).

MOM estimates for differences in life expectancy by education and occupation with Italian MPS data

When the survey data is of good quality and the case numbers are high enough, the MOM provides reliable estimates even for population subgroups. This can be demonstrated with data on paternal orphanhood from the Italian MPS 1998 and 2003 because estimates for the period 1984–1990 can be obtained from both surveys (see Fig. 1). The dot plots in Figures 3 and 4 display the corresponding estimates for life expectancy at age 30 by education level and occupation, respectively. The data from both surveys provide the same order from the lowest to the highest levels of life expectancy for both education and occupation. Even differences between the groups are comparable. The similarity of the results obtained from the two independent cross-sectional surveys indicates that the MOM provides stable results, even for the analysis of population subgroups.

MOM estimates for differences in life expectancy between migrants and non-migrants Germany with data from the German INTSY

A major advantage of the MOM over the traditional variants of the orphanhood method is its flexibility regarding the size of the age intervals. All relationships described in the appendix hold for broader age groups as well what allows producing estimates on the basis of higher case numbers. This has been done for estimating differences in life expectancy between immigrants and natives in Germany on the basis of respondents’ age group 18-30. Figure 5 shows the corresponding estimates for the differences between Turkish and Italian immigrants, respectively, and Germans derived with the MOM for the year 1993 in comparison to the differences between foreigners and Germans according to the data of the 1987 census (note that official German population statistics do not allow any further differentiation). The results fit each other very well. Thus, the MOM supports the census estimates and helps to get more detailed information about the life expectancy of specific immigration groups in Germany.
**Discussion**

The orphanhood method was originally designed to be applied to populations of developing countries. We modified the method for application in populations of developed countries with the main aim to provide a possibility for estimating mortality differentials when no official data are available. The method can be used to reconstruct a time series of female and male survivorship estimates and thus to analyze time trends of mortality differentials from survey information on the proportion of respondents of different ages with mother/father still alive. In this paper we summarized our experiences in applying the MOM to different data and research questions.

The main conceptual difference between the traditional variants of the orphanhood method and the MOM is that the weighting factors and regression coefficients of the traditional methods are based on theoretical population models, independent from the period to which the estimates refer. By contrast, the MOM is based exclusively on empirical data from official statistics for the population analyzed. Another difference is that the traditional method for reference period estimation (Brass and Bamgboye 1981) uses the theoretical time lag between period and cohort-type life tables to display the same survivorship ratio. In the MOM, the reference periods are derived from the average dates of death of the deceased parents.

Although the reported survival of respondents’ parents is cohort survival we followed the approach of the traditional variants of the orphanhood method to estimate past period mortality. Due to the differences between period and cohort-type mortality schedules, the reported cohort survival of respondents’ parents has a different age pattern than the estimated period life table. This difference becomes stronger in ages when the survivorship curve turns downward. Consequently, estimates based on younger ages of respondents are more robust than estimates based on older respondents. According to our empirical applications, age group 60-64 should be the highest to apply the orphanhood method since—in addition to the turn of the survivorship curve—at older ages the case numbers of respondents with parents still alive get too low.

Our tests and empirical applications of the MOM for the indirect estimation of adult mortality in developed populations indicate that the method provides reliable estimates for differences in life expectancy between specific population subgroups. This is important since many developed countries lack official population data to investigate such mortality differences. Adopting results from other countries is an unsatisfying alternative since specific national circumstances are able to strongly influence the magnitude, pattern, and causes of inequalities in health. Moreover, there is an increasing number of high quality survey data with the necessary information about survival and specific characteristics of respondents’ parents, like the Gender and Generations Program (GGP). However, the applicability of
the method is heavily depending on the quality of the survey as demonstrated by our application of the MOM to data of the German GGS. In this respect, analyses based on survey data are always disadvantaged in comparison to direct estimates using large population statistics.

The lower mortality level of a population of parents compared to the total population, as necessarily obtained by orphanhood-based estimates, does not reduce or qualify the applicability of the MOM. It is important to note that the aim of the method is not to provide a new or alternative estimate for the life expectancy of the total population. The aim is to estimate differentials in life expectancy that cannot be analyzed otherwise due to the lack of statistical data. The overestimation of life expectancy can be assumed to affect population subgroups equally which enables a comparison of the differences between them. Murphy et al. (2006) arrived at a similar conclusion regarding their indirect analysis of mortality by education level in Russia based on information about survival of spouses and siblings.

Nevertheless, as Timæus (1991c) has outlined in detail, indirect methods like the orphanhood method always entail several drawbacks. Indirect methods can provide only broad measures of the overall level and trend in adult mortality. They are inherently unable to detect short-term trends or abnormal age patterns of mortality within adulthood. A further limitation of the orphanhood method is that it yields estimates of mortality that refer to dates well before the survey was conducted. Deaths of parents occur over a period extending back to when respondents were born, in the case of mothers and about nine months earlier for fathers. The younger the respondents the more recent are the derived mortality estimates. But even the estimates based on respondents aged 20-24 refer to a period about 8.5 years prior to the survey. This problem could be overcome when two or more surveys containing the necessary information for the orphanhood method and spaced by approximately five or ten years are available so that the ‘cohort-change parental survival methods’ can be applied (see e.g. Hill et al. 1983). Luy (2012) showed that the orphanhood method seems to be a promising estimation technique also in this variant. A specific qualification for estimating mortality differences from maternal and paternal orphanhood as reported in survey data is the appearance of mothers and fathers with unknown characteristics. In our estimation of differences by education and occupation in Italy we found that these parents exhibit a higher mortality than mothers and fathers with available information about socioeconomic status.

On the other hand, the orphanhood method and other indirect methods have some advantages over direct methods. First, they permit the derivation of life tables and thus the estimation of life expectancy. In many cases and for most users of demographic data—like policy-makers—differences in life expectancy are more important and more informative than differences in standardized death rates or relative risks which are usually calculated when surveys or linked census data are used for the analysis of mortality. Years of life represent the most easily understandable unit of measurement of mortality
differentials. Variations in standardized death rates or relative risks are more difficult to assess—above all for laypersons—since large differences in these measures do not necessarily reflect large differences in actual life time. Second, indirect methods typically provide trends of demographic conditions which are derived from a single cross-sectional survey, whereas direct methods usually provide but one estimate from cross-sectional data for a specific year. A third advantage is that the information used is based on respondents’ lifetime experience and thus, fairly precise estimates of the proportions of respondents with living parents (or other relatives) can be obtained even from surveys of moderate size as demonstrated with the German INTSY data. Thus, knowing the general functionality of this method, information on interesting aspects of mortality can be collected quite easily and at a moderate cost by including a few simple questions into existing or planned survey programs. Information about the level and trend of mortality differentials is often sufficient to provide a basis for forecasting and the allocation of resources.

We have to note that results like those presented in this paper call for confidence intervals in principal. Theoretically, there are several possibilities. One is to derive confidence intervals for the proportions of respondents with a mother or father alive by means of bootstrapping or formulae based on asymptotic distributions. The resulting proportions representing the upper and lower limits of the confidence interval can then be used to derive the corresponding survivorship probabilities and estimates for life expectancy. Another alternative is the derivation of confidence bands for the estimated survivorship function, like the ‘Kolmogoroff-Smirnov’ band (based on Kolmogoroff 1941 and Smirnov 1948) or the ‘equal precision band’ (Nair 1984). Usually, these tests were developed to test the statistical significance of differences between complete survival functions (see, for example, Klein and Moeschberger 2005). Nevertheless, the upper and lower bands might also be used to derive the corresponding life expectancies. In our previous applications we refrained from constructing confidence intervals, however. The objective of using the MOM is not to estimate the precise levels of life expectancy for a specific subpopulation, but to obtain information about levels and trends of differences in life expectancy between the subgroups. Thus, the major aim of the MOM does not necessarily require the information whether each difference is statistically significant, above all since the obtained results fit those from other populations with better data.

Finally, the empirical applications of the proposed modification of the orphanhood method provide important information for its traditional variants. The fact that the traditional approaches of Brass and Hill (1973) and Timeæus (1992) and the method of reference date estimation by Brass and Bamgboye (1981) do not perform well in the application to Italian survey data does not mean that they do not work for populations of developing countries. The underlying mortality models of these methods are too different from the mortality of a developed population to be useful here. In developing countries with high mortality levels they should be more appropriate. However, the theoretical considerations
and empirical outcomes regarding the lower mortality levels of parous women and men suggest that these relations hold in every population. According to our knowledge, this effect is not adjusted for in the traditional methods which were developed to obtain estimates of adult mortality for the total population. A further adjustment of the traditional methods to overcome this underestimation of the overall mortality level might be useful.

References


Luy, M. 2009. Estimating mortality differentials in developed populations from survey information on maternal and paternal orphanhood. European Demographic Research Papers, ed. Vienna In-


Figure 1: MOM estimates for life expectancy at age 35, Italian Multipurpose Surveys 1998 and 2003, single estimates for five-year age groups (estimates for 1998 in green, estimates for 2003 in red) and linearly smoothed trends

Source: Luy (2012), own.
Figure 2: MOM estimates for life expectancy at age 30, German Gender and Generations Survey 2005, single estimates for five-year age groups and linearly smoothed trends

(a) Females

(b) Males

Source: own.
Figure 3: MOM estimates for life expectancy at age 30 by education, Italian Multipurpose Surveys 1998 and 2003, males, period 1984-90

(a) MPS 1998

(b) MPS 2003

Figure 4: MOM estimates for life expectancy at age 30 by occupation, Italian Multipurpose Surveys 1998 and 2003, males, period 1984-90

(a) MPS 1998

(b) MPS 2003

Figure 5: Difference in life expectancy at age 30 between foreigners/immigrants and Germans according to the census data of 1987 and MOM estimates for 1993 based on data of the German Integration Survey

(a) Females

(b) Males

Appendix: formal demographic derivation of the MOM

The orphanhood method is based on the relationship between the proportion of respondents with mother/father alive and the demographic experiences of their parents from the moment of respondents’ birth until the time of the survey. The proportion of respondents from an age group (n, n+4) with mother/father still alive is usually denoted by S(n), with n=20 for the age group 20-24, n=25 for the age group 25-29, and so forth. If the time of interview is denoted by T we know that respondents aged (n, n+4) were born on average in T−n, with n being the average age of respondents of that age group. At that time, their mothers/fathers were between α and β years old, with α and β being the minimum and maximum ages at childbearing. Thus, the proportion of respondents whose mother/father survived until time T approximates a weighted average of cohort survivorships

$$S(n) \approx \sum_{\alpha}^{\beta} w_x \cdot \left( \frac{p_{x+\pi}}{p_x} \right)$$

where x represents the single ages at childbearing of respondents’ mothers/fathers, p_x the corresponding cohort life table survival probability to age x, and the weights w_x are the proportions of x-year old mothers/fathers at time T−n, i.e. at the moment of respondents’ birth. These weights are resulting from

$$w_x = \frac{N_x \cdot f_x}{\sum_{\alpha}^{\beta} N_a \cdot f_a}$$

where N_x (respective N_a) and f_x (respective f_a) denote the number of x-year olds (a-year olds) and the fertility rate at age x (age a) at the time of respondents’ birth. Hence, the extent to which the cohorts of parents contribute to the value of S(n) depends on two factors: (i) the age distribution N_x of the parents’ cohorts at the time of respondents’ birth, and (ii) the fertility schedule comprised by the age-specific fertility rates f_x at that time.

The MOM adopts the idea of the classical approaches of the orphanhood method to transfer the mixed cohort survivorship ratios S(n) into period survival probabilities for a specific reference period t(n). We define this reference period as the calendar year in which deceased parents of respondents aged (n, n+4) died on average. Thus
with \(D(n)_y\) being the number of deceased mothers/fathers of respondents aged \((n, n+4)\) during year \(y\) after the birth of respondents, where \(y=1\) for the first year after respondents’ birth, \(y=2\) for the second year and so forth until \(y=\bar{n}\) for the last year of the observed life span of respondents’ parents.

Based on the estimated reference periods \(t(n)\) the empirical values for \(\hat{S}(n)\) are converted into period survivorship probabilities from age 30 to age 33+n. The procedure combines two transformations. The first is the transformation of survival of the cohorts of respondents’ parents from an average age at childbearing \(x\) to \(x + \bar{n}\)—reconstructed from official mortality data with the weights \(\hat{w}_x\) derived from the survey data—into cohort survival from age 30 to 33+n (first fraction on the right hand side of Equation 4), and the second is the transformation of this cohort survival into period survival from age 30 to 33+n for the corresponding reference period \(t(n)\) (second fraction on the right hand side of Equation 4). Thus

\[
\left(\frac{i_{33+n}}{i_{30}}\right)_{t(n)} = \hat{S}(n) \cdot \frac{\sum \hat{w}_x \cdot \left(\frac{p_{33+n}}{p_{30}}\right)}{\sum \hat{w}_x \cdot \left(\frac{p_{x+n}}{p_x}\right)} \cdot \frac{\left(\frac{i_{33+n}}{i_{30}}\right)_{t(n)}}{\left(\frac{i_{30}}{i_{30}}\right)_{t(n)}}
\]

with \(p_x\) and \(l_x\) representing the cohort and period life table survival probability to age \(x\), respectively, and hats indicating that variables refer directly to the parents of survey respondents whereas variables without hats refer to the entire population. Finally, Equation (4) can be simplified to

\[
\left(\frac{i_{33+n}}{i_{30}}\right)_{t(n)} = \hat{S}(n) \cdot \frac{\left(\frac{i_{33+n}}{i_{30}}\right)_{t(n)}}{\sum \hat{w}_x \cdot \left(\frac{p_{x+n}}{p_x}\right)}
\]

More details and numeric examples can be found in Luy (2012).