Multivariate Cohort-Based Analysis of Individual Risk of Mortality in Malawi: The Effects of Children Ever Born and School Attainment

Henry V. Doctor
University of Pennsylvania
Population Studies Center
3718 Locust Walk
Philadelphia, PA 19104
hvdoctor@mail.sas.upenn.edu

Extended Abstract

Introduction

A tremendous body of literature on adult mortality documents that survival is higher among women with less children ever born (CEB) and higher school attainment both in developed and developing countries than those with more CEB and lower school attainment (Kitagawa and Hauser 1973; Okonofua et al. 1992; Lutz, Goujon, and Doblhammer-Reiter 1998; Ross and Mirowsky 1999; Kaye et al. 2003). Thus, CEB is an important predictor of survival due to the negative health effects associated with prolonged childbearing. A number of studies have demonstrated that women with more children are more likely to have negative health outcomes and experience higher mortality than women with less children (Ronsmans 1995; Stanton, Abderrahim, and Hill 2000).

School attainment has positive effects on health because it has a direct effect on individual's income generating ability and on their access to adequate diet, shelter, health care services, and other life-enhancing resources. A number of studies have found that school attainment is closely related to important health behaviors, such as smoking, than is either income or occupation. Schooling also provides opportunities in accessing and using information that may improve one's life chances. Therefore, the well educated experience better health than do the disadvantaged, as indicated by high levels of perceived health and physical functioning and low levels of morbidity and mortality (Ross and Mirowsky 1999).

Despite the extensive body of research in developed and other developing countries that has attempted to assess the pathways by which CEB and schooling operates to affect mortality, little research in Malawi has examined the extent to which the CEB-mortality and schooling-mortality relationship varies across birth cohorts. At least in the Malawian context, a thorough understanding of how CEB and schooling relates to health, however, requires a consideration of cohort effects. Fertility levels have changed substantially in Malawi between the late 1970s and late 1990s. For example, the total fertility rate in 1977 was estimated at 7.6 and declined to 7.4 and 6.2 in 1987 and 1998 respectively (National Statistical Office [Malawi] and ORC Macro 2001). Similarly, school attainment has undergone rapid changes over the twentieth century in Malawi from the period when schooling was introduced by the early missionaries (circa 1860) to the time when the country adopted the multiparty system of government in 1994. These changes make it likely that the relationships among CEB and schooling and its mediators also changed. During the period between late 1970s and late 1990s, the distribution of mortality changed tremendously implying that the effect of CEB and schooling may have changed as a result of changes in the epidemiology of disease.

Because of the changes in the number of CEB and the level of schooling over time, it is reasonable to expect that the effect of CEB and schooling on health and mortality may vary across cohorts. Research of this nature, though fragmented, has been conducted mostly in developed countries. For example, other researchers (e.g., Preston and Elo 1995; Manton et al. 1997) have examined the relationship between schooling and mortality in the United States and have assumed that temporal variations in the relationship is due to cohort or period effects. Studies of this nature are lacking in Malawi.

In this paper, I examine the birth cohort effects that are involved in the relationship between CEB and mortality, before and after adjusting for the level of school attainment. Thus, I investigate the importance of CEB and school attainment in predicting survival of a cohort of women aged 45 and older in the 1987 Malawi census who are followed in the 1998 census. The choice of these variables is based on the requirements of the method employed in this paper which is that all the covariates have to be fixed over time. Specifically, employing schooling has also the advantage in that, unlike some other measures of SES such as occupation and income, it is easily measured and generally fixed for each individual relatively early in adulthood, whereas occupational earnings and status may change in response to changes in health (Freedman and Martin 1999). As a result health problems that emerge among individuals in later life are unlikely to influence the level of school attainment. Moreover, school attainment can be assesed for all individuals, whereas occupational status is not assessed easily for persons who have left or who have never entered the labor force (Christenson and Johnson 1995).

Methodology

To investigate the influence of CEB and school attainment on survival of a cohort of women in one census who are followed into another census, one ideally needs to follow the cohort from whom the number of CEB and school attainment of women are available at baseline and examine the differences in the risk of mortality by birth cohort. To do this, I use an indirect estimation approach which follows three sequential 11-year birth cohorts: 1900-1910, 1911-1921, 1922-1932; and a 10-year birth cohort, 1933-1942, over the 1987-1998 inter-censal period (i.e., a total of four birth cohorts). I infer differential survival across categories of CEB and school attainment by examining changes in the relative risk of survival in each birth cohort.

This study uses age, CEB, and school attainment (no schooling, primary, and secondary+) variables, from the 1987 and 1998 census data for native-born women. Individuals are selected by age; birth year is determined by subtracting age from census year. Individuals are grouped into four birth cohorts based on age reported in the census. The last birth cohort spans to 1942 and its interval is 10 because of the age restriction, that is, those born in 1942 are aged 45 in 1987 whereas those born in 1943 if included, the interval will be 11 and will violate the age limits. This analysis is restricted to women who were aged 45+ in 1987 and were aged 56+ in 1998. These age groups are worth examining because by 1987 individuals aged 45+ had almost completed their fertility and (formal) schooling. This makes the data conform to one of the assumptions of using the method employed in this chapter which is discussed in a later section. In brief, this study uses a method which requires that all explanatory variables be time-invariant and fully comparable for cohort members in the two censuses.

In addition, only persons who are natives are included since it is assumed that the birth cohorts are fixed, with no migration in or out. In-migration is precluded by restricting the analysis to those born in Malawi. I also assume that out-migration is supposedly inconsequentially rare for native-born older women. A cohort enters the analysis when every member is 45 years.

Statistical analysis

Each of the two census micro-data sets is assumed to be independent, representative cross-section of the target population: Malawi-born women with birth years 1900 to 1942. In the analysis I determine whether survival over an 11-year interval between censuses is associated with CEB and school attainment for each birth cohort. I estimate the survival of persons with 1-2, 3-4, … , 9-10, and 11+ children relative to childless women. In addition, I also estimate the association between CEB and survival by controlling for schooling, that is, 1-8 and 12 or more years of schooling, relative to 0 years, for the intercensal period.

Data analysis proceeds by implementing the intercensal survivorship ratio that is used to estimate cohort-specific survival probabilities from two population censuses (Elo and Preston 1994; Hill 1999). When a population is closed and assumed to have perfect reporting, the number of people recorded in the first census represents the total number of people who are at a risk of dying between the first and the second census. It follows that the number of people who survive to the second census are survivors of the original cohort. Let P represent the probability that an individual survives from census 1 (denoted as t_1) to census 2 (denoted as t_2). If N_1 is the number of people alive at t_1 and N_2 is the number of people alive at t_2, then assuming that there is no migration between t_1 and t_2, we can estimate P as the ratio of the second census count to the first census count, that is:

$$P = N_2 / N_1$$

We can therefore estimate census survival probabilities for a closed cohort from two independent samples. Hill (1999) refers to these samples as n_1 and n_2 for the two censuses and presents a cohort-based model of individual survivorship generalized to a multivariate set of continuous and discrete independent variables. The survival process under investigation is assumed to be a single-decrement process in which the only form of leaving the cohort is through "death". The cohort is assumed to be closed to migration, including pseudo-migration caused by age misreporting and other forms of misclassification.

An illustration of Hill's method

Hill (1999) presents the method as follows: Let π_i represent the probability of surviving from time t_1 to time t_2 for individual i where i indexes members of the population at time t_1, for $i = 1, 2, \ldots, N_1$. Assume that this probability depends on a vector of time-invariant covariates $\mathbf{X}_i$ through a log-probability model:

$$\log \pi_i = \alpha + \beta' \mathbf{X}_i \qquad (1)$$

Here, the constant α represents the log of the baseline probability (level) and β_k represents the log of the relative risk (differential) associated with covariate k, adjusted for all other covariates in the model, such that $\exp(\beta_k) = \hat{R}_k$. The term $\hat{R}_k$ is defined as the relative survival probability measuring the association of covariate k with survival (Hill 1999). Assuming that we have a sample of n_1 cohort members (and information about their covariates) at time t_1 and an independent sample of n_2 cohort members (and comparable covariate information) at time t_2, we can let c and d represent the sampling fractions for the t_1 and t_2 samples, respectively. We can further imagine constructing a pooled data set consisting of the two independent samples, and let binary variable y_i equal 1 if observation j is in time t_2 sample and 0 otherwise, for $j = 1,2,\ldots,n_1+n_2$. Finally, let τ_j, represent the probability that observation j is in the time t_2 sample given that it is in the pooled data set. Using the pooled data, one can estimate the parameters of the log probability model in equation (1) by fitting a logit regression model to the dummy indicator y_j:

$$\log \frac{\tau_j}{1-\tau} = \alpha' + \beta' \mathbf{X}_j \qquad (2)$$

where $\alpha' = \alpha + \log(d/c)$. The logit model is justified because the log-odds that an observation in the pooled data set is from t_2 sample equals the log of the survival probability to within an additive constant of log (d/c); see Hill (1999: 498) for details.

Using the 1987 and 1998 Malawi census data I investigate factors influencing survivorship of women by assembling a pooled two-micro data set for the cohort of individuals aged 45+ at the 1987 census and aged 56+ at the 1998 census. The analysis explores the influence of CEB and schooling on survival over the 11-year intercensal period.

I estimate two survivorship models using the logit regression in equation (2). The first model estimates the baseline CEB, unadjusted for the confounding effect of schooling. As discussed earlier, because these women were beyond reproductive age group in 1987, CEB is considered fixed over the period. I hypothesize that women with a high number of CEB will have lower survival chances than those whose number of CEB is low due to the negative health effects associated with higher parity.

The second model adds a trichotomous schooling variable to the logit regression equation since for women who were aged 45+ in 1987, schooling level is considered fixed over the period. In this case, schooling acts as a proxy for SES since the latter is related to survival for women with higher schooling than those with low schooling level. Fertility is also low among women with higher SES.

In a separate analysis, I estimate three survivorship models in which CEB is entered in the first model and the second model controls for schooling level. The last model includes the birth cohort. This analysis is performed in order to compare the results with those in which I estimate three survival models separately for the birth cohort.

Results

Preliminary results show that women with low CEB have higher chances of surviving than those with high CEB. School attainment also matters in predicting survival. Women with primary or secondary schooling have higher chances of surviving than those without. Findings from studies of this nature may influence policy in directing resources towards family planning and schooling programs. These results are also consistent with studies from other developed countries such as the United States that have found that high fertility is associated with negative health effects in women due to strenuous effects of childbearing (see Hill 1999 and Lauderdale 2001). Schooling is the key SES indicator since it is associated with the ability to learn new skills and accumulate human assets. Alongside this understanding, the findings from this study are consistent with the expectation that schooling should have a negative effect on mortality, that is, the highly educated individuals should have higher chances of surviving than those with low levels of education. Because these women were exposed to childbearing and schooling at an earlier age, programs which target women's schooling and family planning should be strengthened at the local and national level.