Effect of Mortality Selection on the Observed Education-Mortality Relationship across a Cohort Life Course

Anna Zajacova[1]
Princeton University

March 2004

[1]Office of Population Research,Princeton University, Princeton NJ 08544.

Abstract

Two opposing hypotheses have been proposed to explain the relationship between education and health across age: cumulative advantage and age-as-leveler. Empirical studies generally found a pattern consistent with the latter perspective, showing a diminishing effect of education on health at older ages. Some researchers suggested that the observed converging lifecourse patterns are an artifact of selective mortality in the population where the true effect is cumulative. The goal of this paper is to show that the observed declining or curvilinear lifecourse pattern between education and mortality could be caused entirely by mortality selection due to unobserved heterogeneity in population frailty. We use simple macrosimulation models to show that in a population characterized by heterogeneity in mortality hazard, a true linearly increasing effect of education on mortality may appear to have the lifecourse pattern observed in many empirical studies.

Socioeconomic status has been shown consistently to be a strong predictor of health and mortality. Understanding the pattern of education-health relationship across the life course is a crucial step in disentangling the pathways that mediate their association. The pathways ultimately operate on the individual psychosocial and biological level, thus we also need to understand the individual-level lifecourse trajectory of the association between education and health (O'Rand 2001). Empirical studies necessarily measure lifecourse patterns at the population level and researchers then make inferences about the individual-level trajectories from the observed population data. These inferences, however, may not straightforward and may potentially lead to erroneous conclusions (Vaupel and Carey 1993).

This paper focuses on exploring the difficulties with inferring individual lifecourse trajectory of education's effect on mortality from population-level data. In particular, I explore the effects of unobserved heterogeneity as a potential cause of bias. Vaupel et al. (1979; 1985a; 1985b) showed how heterogeneous frailty in a cohort biases the shape of the observed mortality hazard. I extend their work to explore the potential bias on estimates of mortality's determinants, such as education, across the lifecourse. Research on determinants of health lacks literature on the effects of unobserved heterogeneity and as a result researchers may make simplifying and potentially misleading inferences about their effect across age (Vaupel, Manton and Stallard 1979; Vaupel and Yashin 1985b).

Lifecourse pattern of SES-health relationship: Theory

There are two major theoretical perspectives in social epidemiology that purport to explain individual lifecourse pattern of the effect of socioecoonmic status on health and mortality. The pattern is a property of a cohort, not of individuals (Dannefer 2003). However, undestanding these cohort patterns will inform individual pathways from SES to health.

The first perspective suggests that the effect of education on health and mortality diminishes in old age. This hypothesis is variously called 'age as leveler' or 'convergence' perspective. According to this theory, the declining effect of education on health and mortality in old age is due to a combination of several factors. Government aid to the elderly in the form of social security and Medicare is thought to balance out some of the socioeconomic inequalities that grew throughout adulthood. Furthermore, health may

become more strongly dependent on age in later years due to a universal biological frailty in late life (House et al. 1994; Markides and Black 1996), making the relative importance of SES predictors smaller.

The second perspective is the cumulative advantage theory, which suggests that the opposite individual-level lifecourse pattern occurs: the effect of socioeconomic status on health and mortality increases with age; that is, health differentials diverge across the lifecourse (Lauderdale 2001; Lynch 2003). According to this perspective, inequality within cohorts increases with age (O'Rand 2001) in a process by which early early socioeconomic disadvantages and advantages cumulate gradually to produce an increasingly heterogeneous cohort (Dannefer 1987). Double jeopardy hypothesis is another closely related concept in health lifecourse research, which says that there is divergence in health trajectories between whites and minorities (Ferraro and Farmer 1996a; 1996b).

Crystal and Shea (1990), and Easterlin et al. (1993) found cumulative advantage process in economic inequality—the effects of social security and other government support to the elderly were outpaced by private income and wealth accumulation. Moreover, many health risks and behaviors, such as smoking or obesity, also accumulate over time and can take decades to affect health (Ferraro and Kelley-Moore 2003). Stress as another factor affecting health that has often been conceptualized as having a cumulative effect on health (Lynch and George 2002; McEwen 1998).

The cumulative advantage perspective finds support in diverse social science disciplines and there is relative consensus that cohorts evidence increasing heterogeneity and inequality with age (O'Rand 1996). The divergence concept has been employed in lifecourse study of subjects ranging widely from widening differences in children's abilities (Scarborough and Parker 2003) to wage inequalities (Bernhardt et al. 2001), to the well-known Matthew effect describing the diverging trajectories of academic careers (Merton and Zuckerman 1968).

Empirical findings of lifecourse patterns of education-mortality relationship

While the cumulative advantage perspective is theoretically well-founded and resonates in many disciplines, empirical research of the SES-health association across lifecourse has not been consistent (Ferraro and Kelley-Moore 2003).

Kitagawa and Hauser's (1973) seminal work on socioeconomic differentials in mortality

in the US established the use of education as a strong predictor of mortality. The authors analyzed mortality ratios by educational level among adults 25-64 and 65+, and found that mortality differences by educational level were considerably smaller for the older group. More recently, in the 1990s, several studies based on data from the National Longitudinal Mortality Study, a large nationally representative survey of the US population, also disaggregated the population by age, and all observed a weaker effect of education on mortality among persons 65+ compared to the working-age group (Backlund, Sorlie and Johnson 1996; Elo and Preston 1996; Preston and Elo 1995; Rogot, Sorlie and Johnson 1992; Sorlie, Backlund and Keller 1995). Somewhat surprisingly, no explanations for the diminishing effect of education were offered in any of these papers, or even of reasoning behind dividing the study population into several age groups.

Additional research on educational differences in mortality, employing data from the US and Europe, also found the same general result of converging educational differentials in health and mortality among the oldest (Christenson and Johnson 1995; Feldman et al. 1989; Kunst and Mackenbach 1994; Mustard et al. 1997). Adler et al. (1993) reviewed further studies and noted that "socioeconomic status differences in health are greatest in middle age and early old age compared with both earlier and later in life" (p. 3141), thus lending support to the convergence hypothesis, albeit also without any further explanations about the causes of this pattern.

The papers reviewed above did not specifically focus on the lifecourse pattern of the education-health relationship. House and colleagues (1990; 1994) were interested explicitly in the age pattern of the SES-health relationship. In both papers they observed an inverse-U shape pattern of the education-health relationship, with diverging health inequalities in the early adulthood and convergence after early old age. They concluded that the pattern represents a true trajectory although they also suggested that selective mortality may contribute to the observed pattern.

In contrast to the large body of papers that found support for the convergence perspective, relatively little support can be found for the cumulative advantage perspective. Ross and Wu (1996) studied how educational attainment effected changes in health in a course of one year. They found that the less-educated evidenced steeper declines than individuals with higher educational attainment, thus lending support to the cumulative advantage hypothesis. The authors did not attempt to reconcile their findings with the

existing body of findings pointing in the opposite direction, and they did not suggest any theoretical explanation for the observed divergence.

Recently, there have been some effort in demography to explore lifecourse process in the education-mortality association. Beckett (2000) attempted to adjust for selective mortality in an analysis of education's effect on health, in the lifecourse of a synthetic cohort. She found converging educational inequalities in health in the old age and concluded that selective mortality does not account for the convergence. However, Lynch (2003) later reanalyzed the same data using hierarchical models and found that selective mortality fully explains the observed convergence in mortality by education in the old age. His analysis determined that, at the individual level, the effect of education on health is cumulative across the lifecourse. A similar conclusion was also reached by Lauderdale (2001) who used an indirect estimation approach with Census data and also found that the effect of education on mortality increases steadily with age.

Research question

Vaupel and Yashin (1985b) suggested that heterogeneity should be considered an explanation when theory and auxiliary individual-level evidence conflict with observed (cohort) lifecourse trajectories. It is, I argue, the case here: the cumulative advantage theory is well-founded, but bulk of the empirical evidence supports the convergence perspective.

My contention is that the weakening effect of education on mortality found in most empirical research may be an artifact of mortality selection due to unobserved heterogeneity. The age trajectory of education's effect on mortality in a *population* is often simply assumed to reflect *individual* lifecourse change. However, the cohort pattern could also be an artifact of selective mortality that induces systematic changes in the composition of the cohort and thus biases empirical findings.

In observational studies, inferences about individual patterns are made from cohort or population-level data. In this simulation paper, I reverse the inference process and study the micro- to macro-level relationship. I will use macrosimulation to show how the cohort lifecourse trajectory of education-mortality relationship can be affected by heterogeneous frailty in the population. Wachter exhorted demographers to make heterogeneity a priority in research (1997). In this paper, I will focus on the ability of unobserved heterogeneity to produce artifactual cohort trajectories, and show that making straightforward inferences

about individual-level lifecourse patterns from population data may be problematic.

Method

The mathematics of heterogeneous frailty and its effect on cohort mortality hazard trajectory

In this section, I will present the basic concepts necessary to study the effects of unobserved heterogeneity in frailty on the shape cohort mortality trajectory. The mathematics were developed by Vaupel and colleagues (Vaupel, Manton and Stallard 1979; Vaupel and Yashin 1985a; Vaupel and Yashin 1985b). I will briefly summarize their work, focusing on results relevant for the present research question. For detailed discussion, proofs and examples, see Vaupel, Manton and Stallard (1979), and Vaupel and Yashin (1985a; 1985b). I will employ the tools introduced in this section to consider what happens to predictors of mortality, such as education, during the lifecourse of a cohort in the presence of unobserved heterogeneity.

Consider a cohort of individuals that are followed from early adulthood. The mortality experience of each individual is captured by the instantaneous probability of dying at age x, $\mu(x)$. Of course, all individuals at a given level of mortality hazard do not have identical lifespans. Age at death thus is a random variable with a probability function described by $\mu(x)$.

Three interrelated functions need to be introduced. The instantaneous probability of dying $\mu(x)$ is defined as negative logarithmic derivative of the survival curve,

$$\mu(x) = -\frac{dlogS(x)}{dx}. \tag{1}$$

Cumulative mortality hazard from birth to age x, denoted $H(x)$ is

$$H(x) = \int_{t=0}^{x} \mu(t)dt. \tag{2}$$

Finally, $S(x)$ represents the likelihood of surviving to age x. Is is a strictly declining function, reflecting the gradual dying out of the cohort, calculated by definition (see 1 above) as:

$$S(x) = e^{-\int_{t=0}^{x} \mu(t)dt} e^{-H(x)} \tag{3}$$

In these three defining equations, age is the only predictor of mortality—all individuals of the same age are implicitly assumed to have the same mortality hazard $\mu(x)$. However,

a wide range of observed and unobserved factors effect mortality hazard. Demographers often use the term frailty to represent the unobserved portion of the variation among individuals of a given age in their likelihood of dying.

Frailty, denoted Z, is a random variable usually defined as a combination of unobserved genetic and nongenetic factors that make an individual's mortality rate $\mu(x)$ systematically different from some other individual's mortality rate (Manton, Stallard and Vaupel 1981). It means simply that individuals at a given age (and a given set of measured predictors) vary in their probability of dying (Vaupel and Yashin 1985a). It is assumed in this paper that the level of frailty Z is constant within individuals across the lifecourse; that is, an individual at any given frailty level at birth remains at that level throughout life. Most research on heterogeneity makes this simplifying assumption, mainly because of mathematical convenience. For an example of an analysis that includes a within-individual time-varying frailty, see Manton (1999). Since age is a random variable, there is still a wide variance in life expectancies for individuals at a given frailty level (Vaupel 1988).

Operationally, the level of frailty for an individual is defined as a proportional change of mortality hazard for an individual, compared to some other individuals' mortality hazard. Let me define an individual with frailty $Z = 1$ as a standard individual whose frailty is denoted $\mu(x, 1)$[1]. By this definition, an individual with frailty level $Z = z$ will have mortality hazard

$$\mu(x, z) = z\mu(x) \tag{4}$$

Cumulative hazard $H(x, z)$ for an individual with frailty level z is computed, from (2) and (4), as

$$H(x, z) = \int_{t=0}^{x} \mu(t, z)dt = \int_{t=0}^{x} z\mu(t)dt = z\int_{t=0}^{x} \mu(t)dt = zH(x) \tag{5}$$

The effect of frailty Z is thus multiplicative both on mortality hazard $\mu(x)$ and cumulative mortality hazard $H(x)$. Finally, the survival curve for an individual with frailty level $Z = z$ is calculated simply as

[1]For simplicity I will in the remainder of the paper drop the specification $Z = 1$ for the standard individual. Thus $\mu(x, 1)$ will be simplified to $\mu(x)$, $S(x, 1)$ to $S(x)$ etc.

$$S(x,z) = e^{-H(x,z)} = e^{-zH(x)} = S(x)^z. \tag{6}$$

For example, if a standard individual has 30% probability of surviving to age x, the likelihood of surviving to that age for an individual at frailty level, say $Z = 2$, will be 0.3^2, or 9%.

Equations (1) to (6) above described *individual* mortality trajectories. However, these are unobservable—for any individual, only age at death is an observable variable, while the entire $\mu(x)$ trajectory is latent. We use information about the ages at death for individuals to get the average mortality rate for a cohort, denoted $\bar{\mu}(x)$, cohort cumulative hazard $\bar{H}(x)$, and cohort survival probability $\bar{S}(x)$[2].

Let me first calculate average mortality $\bar{\mu}(x)$ in a cohort composed of two homogeneous frailty groups. Consider a cohort where one group of individuals has frailty $Z = 1$, and another group $Z = z$. The mortality experience of individuals in group 1 is described by equations (1) to (3), and in group z by equations (4) to (6). Also suppose that at age $x = 0$, both of these groups comprise a fraction of the cohort: $\Pi_1(0) = 1 - \Pi_z(0)$. At age x, the proportion of individuals in each group $\Pi(x)$, will decline proportionally to the survival probability $S(x)$ associated with the respective group: $\Pi_i(x) = \Pi_i(0)S_i(x)$, $i = 1, z$. Because the higher-frailty individuals are removed from the cohort at a faster rate than their lower-frailty counterparts, the fraction of the total surviving cohort comprising the high-frailty group will decline steadily. The cohort mortality hazard always increases more slowly that the individual hazards. Accounting for the changing composition of the cohort, the average mortality rate $\bar{\mu}(x)$ is calculated as

$$\bar{\mu}(x) = \frac{\sum_i \mu(x,i)\Pi(x,i)}{\sum_i \Pi(x,i)}, i = 1, z. \tag{7}$$

As the higher-frailty group gradually comprises ever smaller fraction of the cohort, the average hazard approaches the mortality rate for the lower-frailty group (Keyfitz 1985). Thus the surviving cohort is not a random sample from the cohort at birth but is systematically more robust (Manton, Stallard and Vaupel 1981; Vaupel, Manton and Stallard 1979). Figure 1 below shows two examples of the effect of unmeasured heterogeneity on

[2]See Vaupel, Manton and Stallard (1979), p.453 for proof that the average mortality for the surviving individuals equals the cohort mortality hazard.

the observed mortality hazard trajectory $\bar{\mu}(x)$ relative to the unobserved individual trajectories, in a cohort with two homogeneous frailty groups and different baseline hazards.

The example above represented a cohort with a discrete (Bernoulli) distribution of frailty Z. A more realistic distribution of frailty in a cohort is a continuous distribution of frailty $Z \sim f(z)$. Then $\Pi(x, z)$ in (7) becomes $f_x(z)$ and

$$\bar{\mu}(x) = \frac{\int_z \mu_x(z) f_x(z) dz}{\int_z f_x(z) dz}.$$

By (4), $\mu_x(z)$ in the numerator equals $z\mu(x)$, and from definition of probability density functions, the denominator $\int_z f_x(z)dz = 1$. Hence,

$$\bar{\mu}(x) = \int_z z\mu(x) f_x(z) dz = \mu(x) \int_z z f_x(z) dz = \mu(x)\bar{z}(x), \tag{8}$$

where $\bar{z}(x)$ is the average frailty for the cohort at age x. Equation (8) represents the *fundamental theorem of heterogeneity*, which links individual mortality hazards $\mu(x, z)$ and cohort mortality hazard $\bar{\mu}(x)$ through the distribution of frailty $f(Z)$. It states that the average mortality hazard is proportional to the average frailty in the cohort. As the individuals at higher frailty level—and thus higher mortality hazard— are removed from the cohort, the average frailty level in the cohort $\bar{z}(x)$ declines, and the average mortality hazard $\bar{\mu}(x)$ then declines relative to the standard-individual hazard $\mu(x)$.

Gamma distribution for frailty at birth has been shown to be a particularly analytically tractable distribution for frailty in the study of mortality selection. A random variable Z is gamma distributed if its probability density function is

$$f(Z|\lambda, k) = \frac{\lambda^k}{\Gamma(k)} z^{k-1} e^{-\lambda z}, \text{ for } z > 0,$$

with mean and variance

$$E(Z) = \frac{k}{\lambda}, \text{ and } Var(Z) = \frac{k}{\lambda^2}. \tag{9}$$

Gamma distribution can take different shapes, which depend on the parameter k, called the shape parameter. Gamma distribution with $k = 1$ is the exponential function with parameter λ. Figure 2 shows the probability density functions for $k = 2, 4$ and 8. When $k = 2$, the distribution is positively skewed, describing a cohort where bulk of individuals have similar levels of frailty, with a small number having much higher frailty

levels. As parameter k increases, the distribution becomes similar to the normal, and as $k \to \infty$, variance of the distribution approaches zero—cohort is frailty-homogeneous.

Following Vaupel and Yashin (1985a), I will assume that the mean frailty at birth $\bar{z}(0) = 1$. From (9) this assumptions constrains the parameters λ and k to be equal, $\lambda = k$, and $Var(Z(0)) = 1/k$. It can be shown[3] that when frailty at birth is gamma-distributed with mean= 1, then frailty among individuals surviving to age x is also gamma distributed, with parameters $(\lambda + H(x), k)$. Then by definition (9), the average frailty $\bar{z}$ at age x is

$$\bar{z}(x) = \frac{k}{k + H(x)}. \tag{10}$$

From (10) and (8), we easily compute cohort mortality hazard $\bar{\mu}(x)$. Figure 3 shows two examples of the observed mortality hazard $\bar{\mu}(x)$, compared to the hazard for the standard individual $\mu(x)$, in a cohort with a gamma-distributed frailty and different standard-individual hazard trajectories.

The effects of unobserved heterogeneity on the observed cohort-level effect of education

In the previous analysis, age was the only predictor of mortality hazard, and the aim of the analysis was to determine the relationship between the *observed* mortality hazard and *individual* hazards. In this section I will extend the calculations to consider the effect of unobserved heterogeneity on the estimates of mortality predictors, such as education, across age. After explaining the assumptions, I will show the mathematics for relating the individual and cohort trajectories of the education effect.

Assumptions

I make several assumptions in the analysis. First, like in the previous section, I assume the level of frailty Z to be constant for individuals across age. Second, I begin the simulation at age 25, when I assume all education to be completed. This assumption is rather common in previous research in social epidemiology (for example, see Lauderdale 2001). It does not

[3]See Vaupel, Manton and Stallard 1979, and Vaupel and Yashin 1985a, for detailed discussion and proofs.

mean that the *effect* of education remains constant across the lifecourse—I will explore both converging and diverging education trajectories—but the *level* of achieved education does not change for an individual, similar to the constant within-individual level of frailty. Third, I am making the assumption that education has a causal effect on mortality. Both theoretical perspectives of the lifecourse pattern of education and mortality, convergence and cumulative advantage perspective, make this assumption. Finally, I assume that the level of education and frailty for any individual are not correlated. In terms of the model specification, it means that the distributions of frailty $f(Z)$ and education $f(E)$ in the population are independent of each other.

Calculating the cohort-level education effect

Denote the effect of an additional year of education on mortality for an individual $E(x)$, and the observed effect for a cohort $\bar{E}(x)$. In most empirical studies, the effect of education is assumed to be linear on log mortality. Then comparing the mortality level at any two adjacent levels of education will give us the effect of education $\bar{E}(x)$. Let the two compared groups be represented by subscript δ for individuals with lower education, and subscript 1 for individuals with higher education[4]. A schematic version of the equation estimated in empirical studies is $Y = X\beta + \epsilon$, where X denotes years of education, β denotes the log effect of a unit (say, a year) of education, Y is the change in the log mortality rates per unit change in schooling, and ϵ is the residual variance. So the actual estimated equation is $log\bar{E}(x) = log\bar{\mu}_{\delta}(x) - log\bar{\mu}(x)$. (I am using the bars to emphasize that the empirical studies use only the observed, cohort-level information). Exponentiating this equation gives us the effect of education on mortality as the ratio of observed mortality of the lower-education individuals over higher-education individuals:

$$\bar{E}(x) = \frac{\bar{\mu}_{\delta}(x)}{\bar{\mu}(x)} \tag{11}$$

Empirical studies measure the *cohort* effect of education on mortality $\bar{E}(x)$, and use this variable to infer the *individual-level* effect $E(x)$. In order for this inference to be unbiased, the two effects must be equal $\bar{E}(x) = E(x)$. I will show below that if there is unobserved heterogeneity in the cohort, the two effects will in fact not be equal. I will to

[4]Like in the section above, the subscript 1 will be dropped in subsequent equations

simulate the individual trajectories of the effect of education on mortality $E(x)$, which are not observable in empirical data analysis. Then I will compute the resulting cohort-level trajectories, and compare the two.

In the simulation model, the effect of education is defined the same way as in the empirical data, to be multiplicative on the baseline hazard (note that this equation describes *individual* effects):

$$\mu_\delta(x) = E(x)\mu(x) \tag{12}$$

Clearly, equation (11) above is the same as (12), with the difference that the former describes the the effect of education on mortality hazard for cohort, and the latter for an individual. As I showed in the previous section, for a frailty-homogeneous cohort, these two would be equal. Thus, the observed effect $\bar{E}(x)$ would yield an unbiased estimate of the individual-effect of education $E(x)$.

However, in a cohort with unobserved heterogeneity, both low and high-education groups $\bar{\mu}_\delta(x)$ and $bar\mu(x)$ comprise individuals with different levels of frailty Z, and the observed hazard differs from the individual trajectory as I showed in the preceding section in (8). Because the underlying standard-individual hazards and cumulative hazards in the low and high education groups differ (by the effect of education $E(x)$, from (**??**) clearly the average frailty $\bar{z}(x)$ will change at different rates for the low and high education groups. As a result of this difference, the observed low-education group hazard $\bar{\mu}_\delta(x)$ diverges from the individual-level $\mu_\delta(x)$ at a faster rate than the high-education observed hazard $\bar{\mu}(x)$ diverges from the individual trajectory $\mu(x)$. This disparity causes the observed effect of education to differ from the individual trajectory.

Model parameters

The parameters of the simulation model include the shape of the standard-individual mortality hazard $\mu(x)$, distribution of frailty $f(Z)$, and the effect of education $E(x)$ on the baseline hazard across age. First I will present the 'medium-value' parameters used in the basic model, and I will also vary the parameters for effect of education across age $E(x)$, and the distribution of frailty (the latter by varying parameter k).

The distribution of $\mu(x)$ will be described by Gompertz function:

$$\mu(x) = bc^x \tag{13}$$

The source for parameters b and c is the U.S. life table for 1989-1991 (National Center for Health Statistics 1997). This life tables represents the population mortality trajectory. I used the published values of survival curve $\bar{S}(s)$ to compute the hazard $\bar{\mu}(x)$. I fitted the values of $\bar{\mu}(x)$ in a log-linear model to determine the parameters $\bar{b}$ and $\bar{c}$. Then in the simulation model I chose arbitrary values for the parameters, b^* and c^* and ran the simulation model, noting the resulting cohort values $\bar{b}^*$ and $\bar{c}^*$. I compared those values to the U.S. life table parameters $\bar{b}$ and $\bar{c}$. Then I adjusted the model baseline hazard parameters so that the resulting cohort mortality parameters from the simulation $\bar{b}$ and $\bar{c}$ approach the source parameters, and repeated the process until both parameters matched the source precisely. The resulting parameters used in this paper are

$$b = 0.000385 \text{ and } c = 1.095.$$

The effect of education (no added effect on baseline mortality hazard, or HE as the baseline group) and $LE = 2$ in the basic model. This value is based on estimates by Elo and Preston (1996) and distribution of educational achievement in the US from US Census. Elo and Preston showed that estimates of the effect of an additional year of education centered around 8 percent reduction in log mortality rates in a number of studies - that is, log mortality associated with any year of education is 92% of log mortality associated with one year less education. If the US population is split along the approximate median educational achievement and educational attainment means of each half are computed, there is about 8 years of schooling difference between these two large groups. Another way to think about this difference is that I compare mortality of individuals with grade school education and college diploma—these two groups differ also by 8 years of schooling. The $.92^8 \approx .55$, or the less-educated half suffers about half the mortality, compare to the better-educated half. For easier interpretation of the figures, I chose to use the better-educated group as the baseline and the less-educated group will then suffer mortality approximately $\frac{1}{0.5} \approx 2$ times higher.

Results

Figure 4 below shows the main results of the simulation model. It contains a set of nine plots that vary in the parameters $E(x)$ and k. All plots shows the individual and cohort lifecourse trajectories in the effect of education on mortality. The individual trajectories are stipulated by the simulation model. The observed trajectories result from observing the aging frailty-heterogeneous cohort and capturing the resulting trajectories. Thus the simulation model goes from individual-level to cohort-level pattern. In an observational study, we would not have access to the individual trajectories, only the observed cohort patterns. Having the both together will allow us to map the cohort patterns and individual patterns together.

The rows of the plots in figure 4 show different lifecourse trajectories of the individual-level effect of education. In the first row, the effect of education on mortality is unchanging with age. The effect is 8% lower mortality hazard for each year of additional schooling, as suggested by Elo and Preston (1996). In the second row, the effect of education is increasing through age, as the cumulative-advantage perspective suggests. Finally, in the third row, the individual trajectory shows the weakening effect of education — describing the age-as-leveler perspective. In the latter two cases, the average effect across age remains 8%. The columns vary the distribution of frailty in the cohort. The first column shows a cohort with a distribution defined by $k = 2$, the middle column $k = 4$, and the third column the distribution is defined by $k = 8$.

The first three plots show that if the individual trajectory was flat, that is, if the effect of education remained constant across age, we would observe a converging pattern at the cohort level. The bias in the cohort effect, compared to the individual-level effect, increases with age. In young adulthood, mortality hazard is very low, so there is little mortality selection affecting the results. At the hazard increases exponentially, by middle adulthood, the mortality selection begins to start taking the high-frailty individuals out of the cohort at higher rates, thus creating the bias. The higher the variance of the frailty in cohort (associated with lower parameter k), the larger degree of bias occurs. However, even in a cohort with relatively little variance in frailty as shown in third column, the observed effect is converging in the old age.

The second row of plots describes individual-effect of education that correspond to the cumulative advantage perspective. This series shows that even under various spec-

ifications of frailty, we would not observe the diverging pattern - there may be some cumulative advantage in up to middle adulthood but in the old age the cohort data would show convergence. That pattern is similar to inverse U-shape finding by House et al. (1990; 1004). This group of plots shows that unobserved heterogeneity has impact on the observed patterns that is strong enough to completely reverse the general lifecourse pattern observed for the cohort.

Finally, the third row of plots displays individual lifecourse trajectories corresponding to the age-as-leveler or convergence perspective. The resulting cohort pattern is also convergence, albeit with a more steep pattern of weakening of the effect of education. All these nine plots show that the cohort effect is biased downward - that is, compared to the individual-level effect, the cohort effect of education is *weaker*, and the difference increases with age.

In addition, I have also explored the inverse U shape of the effect of education on mortality that House et al. (1990, 1994) found. Figure 5 below shows the individual and cohort-level effects of education when the individual trajectory is quadratic. As I discussed above, the cohort pattern first closely follows the individual pattern as mortality selection is weak in young adulthood. As the aging cohort experiences higher mortality overall, the observed effect becomes more strongly downward biased. However, as figure 5 shows, if the underlying individual-lever trajectory was indeed quadratic, we would observe a similar quadratic pattern in the cohort.

Mathematically, we can show in general that the difference between the individual and cohort level trajectories is a necessary consequence of unobserved heterogeneity in a cohort. [5] Equation (11) describes the relationship between the group mortality rates and the cohort effect of education $\bar{E}(x) = \bar{\mu}_\delta(x)/\bar{\mu}(x)$. We also know from (8) that $\bar{\mu}(x) = \mu(x)\bar{z}(x)$, and finally from (12) we can substitute $\mu_\delta(x) = E(x)\mu(x)$.

Thus,

$$\bar{E}(x) = \frac{\bar{\mu}_\delta(x)}{\bar{\mu}(x)} = \frac{\mu_\delta(x)\bar{z}_\delta(x)}{\mu(x)\bar{z}(x)} = \frac{E(x)\mu(x)\bar{z}_\delta(x)}{\mu(x)\bar{z}(x)}.$$

In cohort with gamma-distributed frailty, $\bar{z}(x)$ was calculated from k and the cumulative hazard for the standard individual $H(x)$ as $\bar{z}(x) = k/[k + H(x)]$. Above I showed (using the effect of frailty z in (5) that an effect that is multiplicative on the mortality hazard

[5] This calculation depends on gamma-distributed frailty, but other, less tractable distributions, which could be modeled by microsimulation, would yield comparable results.

is also multiplicative on the cumulative mortality hazard. This applies to the effect of education $E(x)$ (see (12)), so that $H_\delta(x) = E(x)H(x)$. Then

$$\bar{E}(x) = \frac{E(x)\mu(x)\bar{z}_\delta(x)}{\mu(x)\bar{z}(x)} = \frac{E(x)\bar{z}_\delta(x)}{\bar{z}(x)} = \frac{E(x)[k/k + E(x)H(x)]}{k/k + H(x)}.$$

With some simple algebra, we finally obtain

$$\bar{E}(x) = \frac{k + H(x)}{k/E(x) + H(x)}. \tag{14}$$

This is a general expression that relates the *individual* effect of education on mortality and the *cohort* effect through distributions of frailty Z and the baseline mortality hazard $\mu(x)$. This expression depends on the distribution of frailty specified by gamma distribution, but it does not depend on the shape of the baseline mortality hazard. That is, even if the baseline hazard was constant (or had some other shape, not Gompertz as I used in the simulation), we would observe the same general pattern of the divergence between individual and cohort patterns in the effect of education.

Clearly, as $k \to \infty$, the cohort observed $\bar{E}(x)$ will equal the individual trajectory $E(x)$. Since $\sigma^2 = 1/k$, such a distribution describes a cohort with no variance—hence, a homogeneous cohort. This proves that a homogeneous cohort will show no disparity between the observed cohort effect of education $\bar{E}(x)$ and the individual trajectories $E(x)$.

Also, equation (14) above shows that for a given distribution of frailty in population and individual-level effect of education, k and $E(x)$, as the cumulative hazard $H(x)$ necessarily increases with age, the ratio will gradually approach 1. This shows that the bias in the observed effect of education increases with age, which all figures displayed graphically. Furthermore, if the individual effect of education $E(x) = 1$, that is, if there is no effect of education on mortality hazard, then there is also no bias in the cohort effect of education $\bar{E}(x) = E(x)$.

Discussion and summary

Research on unobserved heterogeneity has been employed by demographers to explain the deceleration of mortality hazard growth at the oldest ages in humans and other species (Vaupel 1997), as well as mortality crossovers at the oldest ages that have been observed for pairs of groups in a population (Johnson 2000; Wing et al. 1985). In this paper, I extended the analysis of the effect of unobserved heterogeneity to study its effect on the coefficients for education as a predictor of mortality across age.

Using mathematical simulation, I showed that unobserved heterogeneity in a cohort systematically biases the coefficient for education as a predictor of mortality toward zero. The analysis does not allow us to prove or reject either the convergence or cumulative divergence perspectives on lifecourse trajectory of the effect of education. What the analysis did show is that if the convergence theory is correct, then we would observed cohort pattern that is generally comparable to the individual trajectory. However, I also showed that if the divergence theory is correct, that is, if the individual lifecourse pattern for the effect of education is difergence, it is possible for the observed pattern to have instead the opposite, convergence or inverse U-shape pattern.

Thus I was able to reconcile the cumulative advantage theory with evidence that mostly supports the convergence theory. Although the degree of bias depends on the value of the parameters in the model, the basic result is very robust: unless the cohort is homogeneous (all heterogeneity is captured by the predictors in the model), which is not very likely, the observed cohort-level lifecourse patterns will differ from the individual-level trajectories. Using different parameters in the simulation model would certainly effect the degree of bias in the observed pattern, compared to the individual trajectory of the education effect on mortality, but the general pattern of findings is very likely to remain unaffected.

Unobserved heterogeneity in a model is basically problem of omitted variables. More specifically, unobserved heterogeneity can be compared to the Simpson's paradox (Pearl 2000) in which compositional differences within groups produce counterintuitive and biased results for the entire population. In cross-sectional models, unobserved heterogeneity put in the error term does not bias the estimates of coefficients if it is not correlated with the predictors. In study of lifecourse patterns, however, unobserved heterogeneity that has effect on mortality will bias the estimates of predictors even if they are not correlated.

There is a limited literature on the effects of unobserved heterogeneity in survival anal-

ysis models (Heckman and Singer 1984; Heckman and Singer 1982; Manton, Stallard and Vaupel 1981; Trussell and Richards 1985). Most research on SES-health relationship, however, neglects the presence of heterogeneity in the population. Social science is ultimately interested in understanding individual-level process, not just describing population-level phenomena. However, researchers need to be aware of the problematic issues involved in making individual inferences from population data gathered from observational studies.

References

Adler, N.E., W.T. Boyce, M.A. Chesney, S. Folkman, and s.L. Syme. 1993. "Socioeconomic Inequalities in Health. No Easy Solution." JAMA 269(24):3140-3145.

Backlund, E., P.D. Sorlie, and N.J. Johnson. 1996. "The Shape of the Relationship between Income and Mortality in the United States. Evidence from the National Longitudinal Mortality Study." Annals of Epidemiology 6(1):12-20.

Beckett, M. 2000. "Converging Health Inequalities in Later Life – an Artifact of Mortality Selection?" Journal of Health and Social Behavior 41(1):106-119.

Bernhardt, A., M. Morris, M.S. Handcock, and M.A. Scott. 2001. Divergent Paths: Economic Mobility in the New American Labor Market. New York: Russell Sage Foundation.

Christenson, B.A.and N.E. Johnson. 1995. "Educational Inequality in Adult Mortality: An Assessment with Death Certificate Data from Michigan." Demography 32(2):215-229.

Crystal, S.and D. Shea. 1990. "Cumulative Advantage, Cumulative Disadvantage, and Inequality among Elderly People." Gerontologist 30(4):437-443.

Dannefer, D. 1987. "Aging as Intracohort Differentiation: Accentuation, the Matthew Effect, and the Life Course." Sociological Forum 2:211-236.

—. 2003. "Cumulative Advantage/Disadvantage and the Life Course: Cross-Fertilizing Age and Social Science Theory." The Journal of Gerontology: Psychological and Social Sciences 58(6):S327-S337.

Easterlin, R.A., D.J. Macunovich, and E.M. Crimmins. 1993. "Economic Status of the Young and the Old in the Working-Age Population, 1964 and 1987." in The Changing Contract Across Generations, edited by V.L. Bengtson and W.A. Achenbaum. Newbury Park, CA: Sage.

Elo, I.T.and S.H. Preston. 1996. "Educational Differentials in Mortality: United States, 1979-85." Social Science and Medicine 42(1):47-57.

Feldman, J.J., D.M. Makuc, J.C. Kleinman, and J. Cornoni-Huntley. 1989. "National Trends in Educational Differentials in Mortality." American Journal of Epidemiology 129(5):919-933.

Ferraro, K.F.and M.M. Farmer. 1996a. "Double jeopardy to health hypothesis for African Americans: analysis and critique." Journal of Health and Social Behavior 37(1):27-43.

–. 1996b. "Double Jeopardy, Aging as Leveler, or Persistent Health Inequality? A Longitudinal Analysis of White and Black Americans." Journal of Gerontology: Social Sciences

51B(6):S319-328.

Ferraro, K.F.and J.A. Kelley-Moore. 2003. "Cumulative Disadvantage and Health: Long-Term Consequences of Obesity?" American Sociological Review 68(5):707-729.

Heckman, J.and B.H. Singer. 1982. "Population Heterogeneity in Demographic Models." in Multidimensional Mathematical Demography, edited by K.C. Land and A. Rogers. New York: Academic Press.

Heckman, J.J.and B.H. Singer. 1984. "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data." Econometrica 52(2):271-320.

House, J.S., R.C. Kessler, A.R. Herzog, R.P. Mero, A.M. Kinney, and M.J. Breslow. 1990. "Age, Socioeconomic Status, and Health." The Milbank Quarterly 68(3):383-411.

House, J.S., J.M. Lepkowski, A.M. Kinney, R.P. Mero, R.C. Kessler, and A.R. Herzog. 1994. "The Social Stratification of Aging and Health." Journal of Health and Social Behavior 35(3):213-234.

Johnson, N.E. 2000. "The Racial Crossover in Comorbidity, Disability, and Mortality." Demography 37(3):267-283.

Keyfitz, N. 1985. "Heterogeneity and Selection in Population Analysis." in Applied Mathematical Demography. New York: Springer-Verlag.

Kitagawa, E.M.and P.M. Hauser. 1973. Differential Mortality in the United States: a Study in Socioeconomic Epidemiology. Cambridge, MA: Harvard University Press.

Kunst, A.E.and J.P. Mackenbach. 1994. "The Size of Mortality Differences Associated with Educational Level in Nine Industrialized Countries." American Journal of Public Health 84(6):932-937.

Lauderdale, D.S. 2001. "Education and Survival: Birth Cohort, Period, and Age Effects." Demography 38(4):551-561.

Lynch, S.M. 2003. "Cohort and Life-Course Patterns in the Relationship between Education and Health: a Hierarchical Approach." Demography 40(2):309-331.

Lynch, Scott Michael, and Linda K. George. 2002. "Interlocking Trajectories of Loss-Related Events and Depressive Symptoms among Elders." Journal of Gerontology: Social Sciences 57B:S117-S125.

Manton, K.G. 1999. "Dynamic Paradigms for Human Mortality and Aging." Journal of Gerontology: Biological Sciences 54(6):B247-254.

Manton, K.G., E. Stallard, and J.W. Vaupel. 1981. "Methods for Comparing the Mortality

Experience of Heterogeneous Populations." Demography 18(3):389-410.

Manton, K.G.and A.I. Yashin. 1999. "Inequalities of Life: Statistical Analysis and Modeling Perspectives." in Age Between Nature and Culture, edited by C. Sauvain-Dugerdil. Geneva, Switzerland: University of Geneva.

Markides, K.S.and S.A. Black. 1996. "Race, Ethnicity, and Aging: The Impact of Inequality." Pp. 153-171 in Handbook of Aging and the Social Sciences, edited by R.H. Binstock and S.A. Black. New York: Academic Press.

McEwen, Bruce S. 1998. "Stress, Adaptation, and Disease: Allostasis and Allostatic Load." Annals of the New York Academy of Sciences 840:33-44.

Merton, R.K. 1995. "The Thomas Theorem and the Matthew Effect." Social Forces 74(2):379-423. Merton, R.K.and H.A. Zuckerman. 1968. "The Matthew Effect in Science: the Reward and Communication Systems of Science are Considered." Science 199(3810):55-63.

Mustard, C.A., S. Derksen, J.M. Berthelot, M. Wolfson, and L.L. Roos. 1997. "Age-Specific Education and Income Gradients in Morbidity and Mortality in a Canadian Province." Social Science and Medicine 45(3):383-397.

National Center for Health Statistics. 1997. "U.S. Decennial Life Tables for 1989-91. vol.1, no.1." Hyattsville, MD.

O'Rand, A.M. 1996. "The Precious and the Precocious: Understanding Cumulative Disadvantage and Cumulative Advantage over the Life Course." The Gerontologist 36(2):239-239.

–. 2001. "Stratification and the Life Course: The Forms of Life Course Capital and Their Interrelationships." in The Handbook of Aging and the Social Sciences, edited by R.H. Binstock and L.K. George. San Diego: Academic Press.

Pearl, J. 2000. Causality: Models, Reasoning, and Inference. Cambridge: Cambridge University Press.

Preston, S.H.and I.T. Elo. 1995. "Are Educational Differentials in Adult Mortality Increasing in the United States?" Journal of Aging and Health 7(4):476-496.

Rogot, E., P.D. Sorlie, and N.J. Johnson. 1992. "Life Expectancy by Employment Status, Income, and Education in the National Longitudinal Mortality Study." Public Health Reports 107(4):457-461.

Ross, C.E.and C.L. Wu. 1996. "Education, Age, and the Cumulative Advantage in Health." Journal of Health and Social Behavior 37(1):104-120.

Scarborough, H.S.and J.D. Parker. 2003. "Matthew Effect in Children with Learning Disabilities: Development of Reading, IQ, and Psychosocial Problems from Grade 2 to Grade 6." Annals of Dyslexia 53:47-71.

Sorlie, P.D., E. Backlund, and J.B. Keller. 1995. "US Mortality by Economic, Demographic, and Social Characteristics: the National Longitudinal Mortality Study." American Journal of Public Health 85(7):949-956.

Trussell, J.and T. Richards. 1985. "Correcting for Unmeasured Heterogeneity in Hazard Models Using the Heckman-Singer Procedure." Pp. 242-276 in Sociological Methodology, edited by N.B. Tuma. San Francisco: Jossey-Bass.

Vaupel, J.W. 1988. "Inherited Frailty and Longevity." Demography 25(2):277-287. . 1997. "Trajectories of Mortality at Advanced Ages." Pp. 17-37 in Between Zeus and the Salmon, edited by K.W. Wachter and C.E. Finch. Washington, DC: National Academy Press.

Vaupel, J.W.and J.R. Carey. 1993. "Compositional Interpretations of Medfly Mortality." Science 260(5114):1666-1667.

Vaupel, J.W., K.G. Manton, and E. Stallard. 1979. "The Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality." Demography 16(3):439-454.

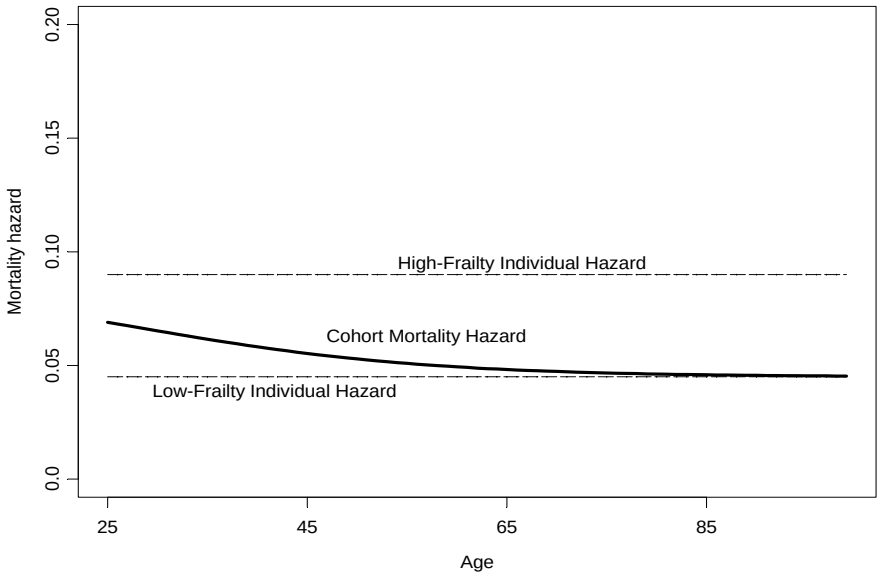
Vaupel, J.W.and A.I. Yashin. 1985a. "The Deviant Dynamics of Death in Heterogenous Populations." Pp. 179-211 in Sociological Methodology, edited by N.B. Tuma. San Francisco: Jossey-Bass.

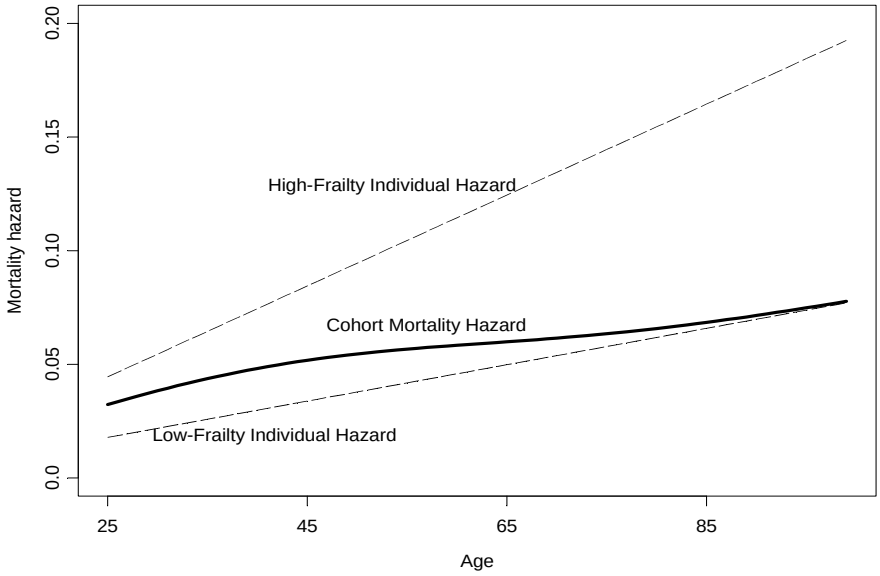
–. 1985b. "Heterogeneity's Ruses: Some Surprising Effects of Selection on Population Dynamics." The American Statistician 39(3):176-185.

Wachter, K.W. 1997. "Introduction." Pp. 1-16 in Between Zeus and the Salmon, edited by K.W. Wachter and C.E. Finch. Washington, DC: National Academy Press.

Wing, S., K.G. Manton, E. Stallard, C.G. Hames, and H.A. Tryoler. 1985. "The Black/White Mortality Crossover: Investigation in a Community-Based Study." Journal of Gerontology 40(1):78-84.

Figure 1: Individual and Cohort Mortality Hazards in a Cohort with 2 Homogeneous-Frailty Groups





Note: In the first plot, frailty is distributed as $z = 1$ and $z = 2$, and $\mu(x) = 0.08$. In the second plot, the two mortality hazards are specified separately.

Figure 2: Gamma Density function with parameter k=2,4 and 8

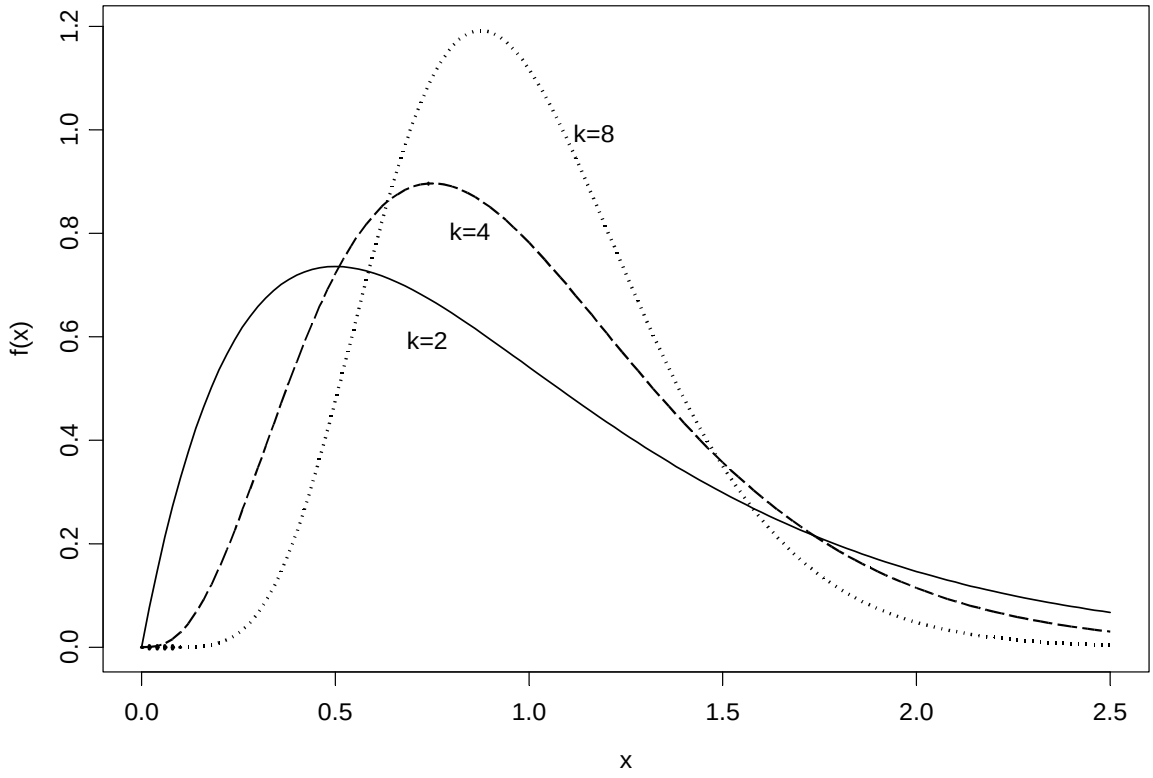
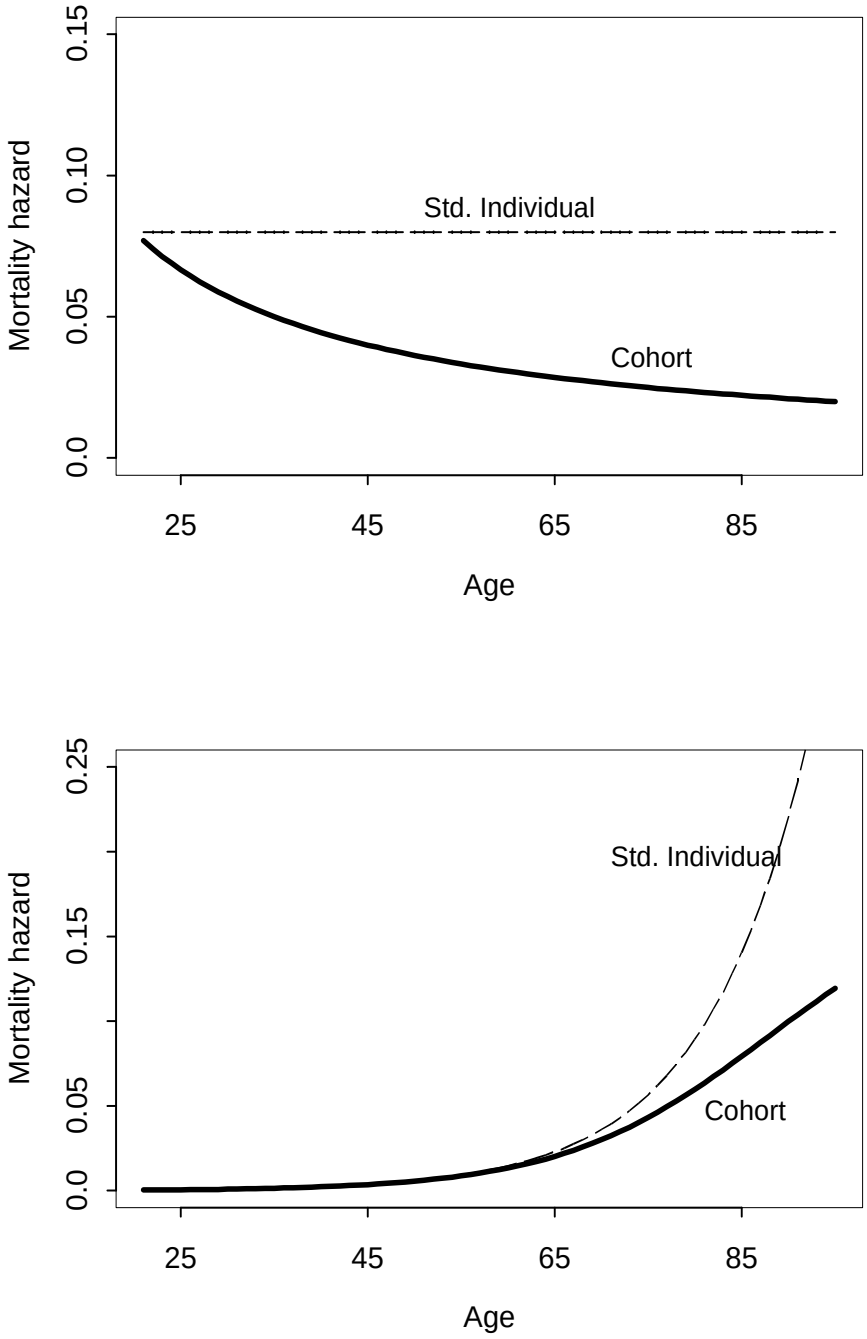
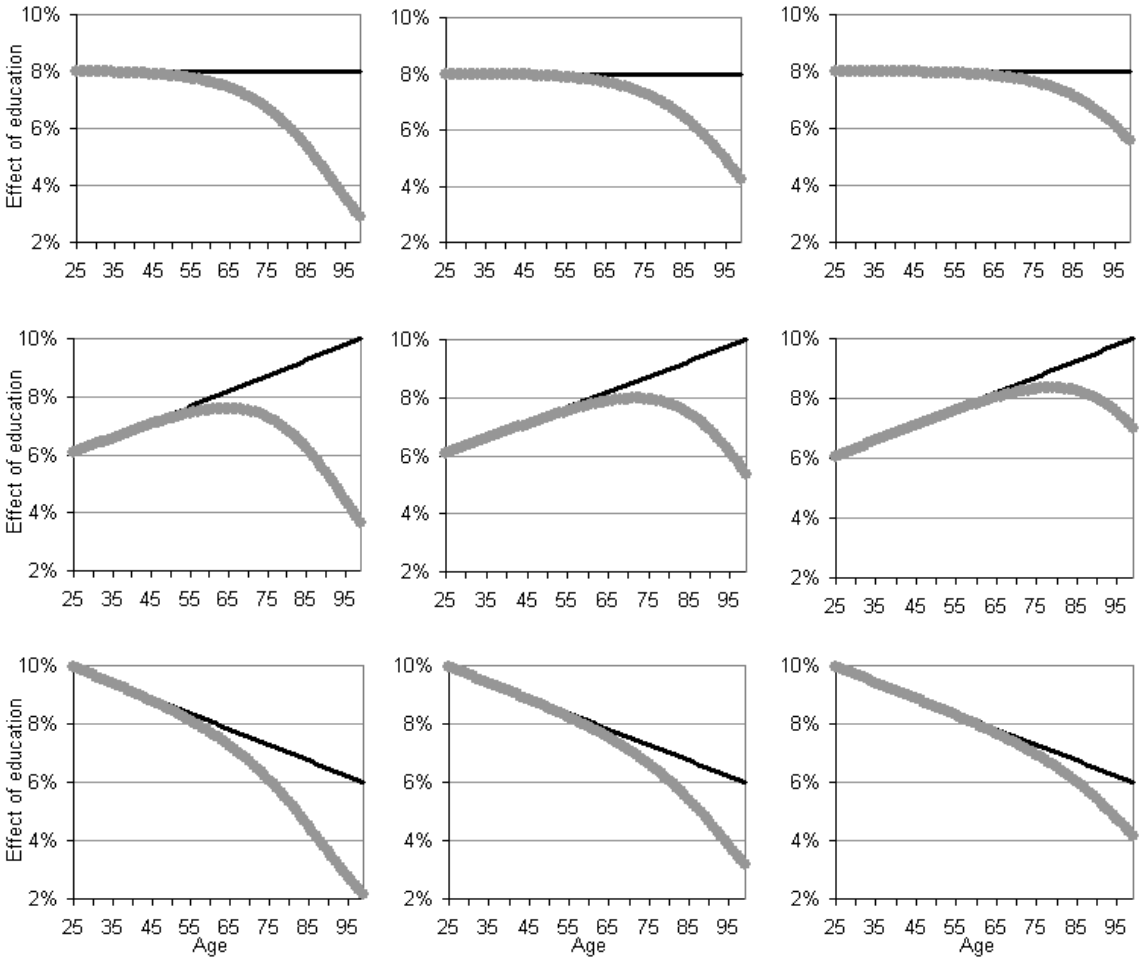


Figure 3: Individual and Cohort Mortality Hazards in a Cohort with Continuously Distributed Frailty.



Note: In the both plots, frailty is gamma distributed with parameter $k = 2$. In the first plot, $\mu(x) = 0.08$ and in the second $\mu(x)$ is Gompertz-distributed with parameters(**??**).

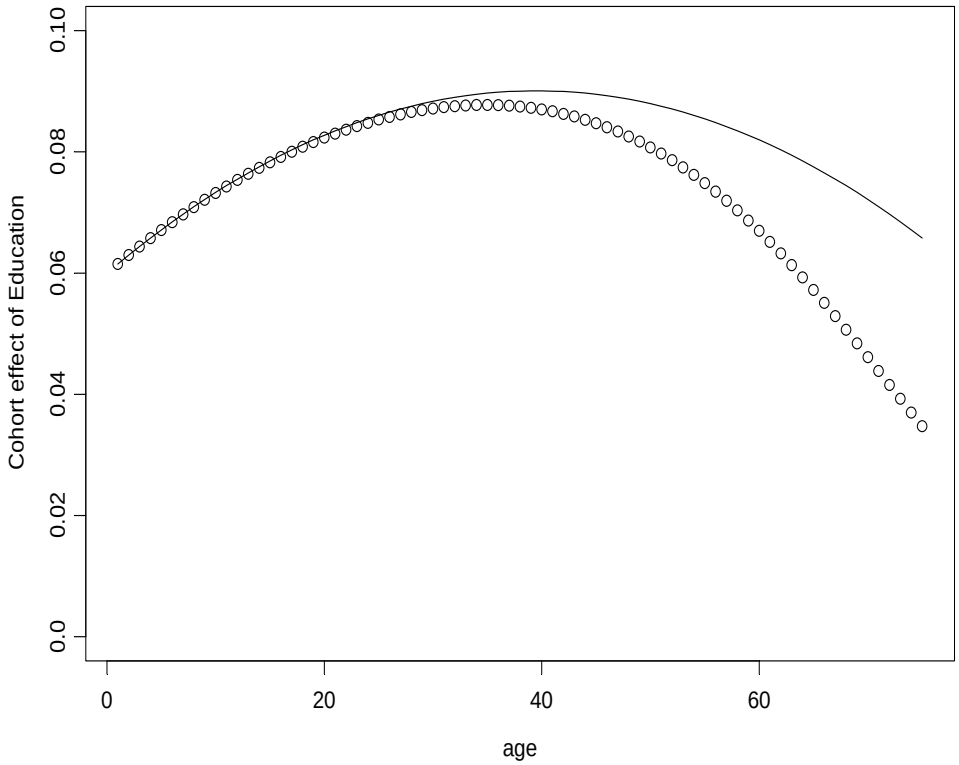
Figure 4: Individual and Cohort Mortality Hazards in a Cohort with Continuously Distributed Frailty.



inin

Note: In the first column, $k = 2$, second column $k = 4$, and in the third column $k = 8$. Each row contains a different lifecourse trajectory of the effect of education on mortality. In all plots, mortality hazard $\mu(x)$ is Gompetrz-distributed with parameters as in (**??**), frailty is gamma distributed with parameter $k = 4$.

Figure 5: Effect of Education—Individual and Cohort Trajectories.



Note: The individual-level effect of education $E(x)$ is defined by $.000019277x^2 - .001522887x + 0.94$. Frailty is gamma distributed with parameter $k = 4$, mortality hazard $\mu(x)$ is Gompertz-distributed with parameters (**??**).