Demography (2016) 53:1135–1168
DOI 10.1007/s13524-016-0489-3

The Effect of Schooling on Mortality: New Evidence From 50,000 Swedish Twins

Petter Lundborg[1] · **Carl Hampus Lyttkens**[2] · **Paul Nystedt**[3]

Published online: 8 July 2016

Abstract By using historical data on about 50,000 twins born in Sweden during 1886–1958, we demonstrate a positive and statistically significant relationship between years of schooling and longevity. This relation remains almost unchanged when exploiting a twin fixed-effects design to control for the influence of genetics and shared family background. This result is robust to controlling for within-twin-pair differences in early-life health and cognitive ability, as proxied by birth weight and height, as well as to restricting the sample to MZ twins. The relationship is fairly constant over time but becomes weaker with age. Literally, our results suggest that compared with low levels of schooling (less than 10 years), high levels of schooling (at least 13 years of schooling) are associated with about three years longer life expectancy at age 60 for the considered birth cohorts. The real societal value of schooling may hence extend beyond pure labor market and economic growth returns. From a policy perspective, schooling may therefore be a vehicle for improving longevity and health, as well as equality along these dimensions.

Keywords Mortality · Longevity · Schooling · Stratified partial likelihood · Twins

Introduction

Longevity is a core component of human welfare, and a wide body of research has established that, on average, people with greater schooling live longer (see, e.g., Adams

✉ Petter Lundborg
petter.lundborg@nek.lu.se

[1] Centre for Economic Demography, IZA, HEP, Lund University, Lund, Sweden

[2] Lund University, HEP Lund, Lund, Sweden

[3] Jönköping Academy, Jönköping International Business School, Jönköping University, Jönköping, Sweden

et al. 2003; Cutler et al. 2006; Cutler and Lleras-Muney 2012). Despite hundreds of empirical studies on the topic, much less is known about why this is so. Rather than a causal relationship, the relationship between schooling and mortality could reflect the influence of unobserved factors affecting both schooling and mortality, such as innate endowments or environmental conditions during childhood (for reviews of research studying the influence of early-life conditions on adult outcomes, see Almond and Currie 2011; Currie 2009). Without a proper understanding about the relationship, it is difficult to assess the value of policies that aim to increase life expectancy and reduce socioeconomic disparities in longevity by investing in greater schooling.

In this article, we contribute to the scant literature on the effect of schooling on mortality. We exploit a unique, historical data set on 50,000 monozygotic (MZ) and dizygotic twins (DZ) who were born in Sweden between 1886 and 1958. Through personal identifiers, we link these data to the administrative data from the Swedish national mortality register, enabling us to observe complete life spans at the individual level for a large fraction of the sample. In addition, through a variety of administrative sources, we construct individual-level data on educational attainment for almost the entire sample. With these data, we can follow individuals through a large part of their life course and study the effect of schooling on mortality at different ages and for different cohorts.

To rule out the influence of unobserved underlying family-level factors influencing both schooling and mortality, we rely on a twin fixed-effects design: we relate differences in schooling between twins to differences in their mortality. Although the twin fixed-effects design relies on strong assumptions, our data allow us to address some of the commonly suggested threats to the design (for a critique of the twin design, see Bound and Solon 1999). MZ twins, who share DNA, may face different environments in the womb that may generate differences in traits such as birth weight. Because birth weight has been found to affect later schooling, twin fixed-effects estimates of the relationship between schooling and mortality might reflect the influence of environmentally induced birth weight differences if these are not controlled for. We can test for this, however, because we can link to unique historical birth records on birth weight for a large subsample of the twins. These records originate from local delivery archives around Sweden that were digitalized by the Swedish Twin Registry. The data cover more or less all twins born between 1926 and 1958.

In addition, we can account for some of the influence of postnatal health and cognitive ability differences between twins. Because a more able twin may invest more in both schooling and health, this may explain part of the positive relationship between schooling and longevity. To address this, we use information on height, a variable known to be correlated with both childhood IQ and early-life health (Case and Paxson 2008; Lundborg et al. 2014; Persico et al. 2004). We link the data on height from the twin surveys. An absence of a change in results when accounting for this trait is informative: it would suggest that neither unobserved ability nor early-life health greatly influences our estimates. In the same spirit, we run regressions on only twins that are very similar in terms of birth weight and height.

Our article relates to the literature seeking to establish causal interpretations of the association between schooling on adult mortality. The strongest evidence comes from a handful of studies exploiting reforms in the education sector. Several of these studies have relied on measures of mortality at the cohort level rather than on individual data.

For instance, Lleras-Muney (2005) exploited changes in compulsory schooling policies in the United States and child labor laws at the state level across the United States and showed that an extra year of schooling reduced the 10-year mortality rate by more than 30 %. Mazumder (2008) showed that the results were not robust to the inclusion of the state-specific trends (although including them risks that they will pick up the effect that one is trying to estimate). Using mortality data at the cohort level, Clark and Royer (2012) exploited two British schooling reforms in combination with a regression discontinuity design but found no effect of schooling on mortality.

A few studies using instrumental variables (IV) have relied on individual-level data. Albouy and Lequien (2009) exploited two compulsory schooling reforms in France and found no effect of schooling on mortality, but their study suffered from small sample size and weak first-stage regressions. Using differences across cohort in the exposure to a Dutch schooling reform, van Kippersluis et al. (2011) found a positive effect of schooling on mortality. Their data included only individuals surviving to the age of 80, however. Using a Swedish schooling reform, Lager and Torssander (2012) provided intention-to-treat estimates of reform exposure and found some evidence for lower mortality in the exposed group. Because of the design of the reform, only a small portion of the population was affected by the reform, however. Similarly, Lleras-Muney (2005) found that only 5 % of the cohort used in the analysis was affected by the reforms. Gathmann et al. (2015) used data from 18 European countries (not including Sweden) and estimated the effect of exposure to schooling reforms on mortality. They found small gains for men but no gains for women.

To the best of our knowledge, the only previous twin-based study on the relationship between schooling and mortality is that by Behrman et al. (2011). They examined the effect of schooling on hospitalizations and mortality among Danish twins born during 1921–1950. Their mortality outcome variable was whether a twin died before 2003. They found that the negative association between schooling, on the one hand, and hospitalizations and mortality, on the other hand, disappeared when they used a twin fixed-effects design. Unlike Behrman et al. (2011), we use a longer period, use a larger sample, and test for the influence of a number of factors that may vary within twin pairs, such as birth weight and height. In addition, we use information on exact death dates in our analysis.[1]

To sum up, the limited literature warrants more empirical work on the effect of schooling on mortality. The results from the few existing studies are mixed, and only a few studies suggest a protective effect of schooling. We contribute to the literature by exploiting unique historical data on schooling and mortality for a large number of individuals from many cohorts. In contrast to the aforementioned studies, we follow a large fraction of individuals throughout the life cycle. Moreover, with the twin fixed-effects design, we can exploit variation across the entire schooling distribution, whereas most IV-based studies exploit variation induced by reforms affecting mainly the lower end of the schooling distribution (for an exception, see Clark and Royer 2012).

Our results suggest that schooling does matter for longevity. We find a strong and positive association between schooling and longevity among both males and females.

[1] More studies have used the twin fixed-effects design to study the effect of schooling on various health outcomes, such as self-reported health, smoking, and obesity (e.g., Amin et al. 2013, 2015a; Behrman et al. 2015; Fujiwara and Kawachi 2009; Lundborg 2013; Madsen et al. 2014; Webbink et al. 2010).

When we apply the twin fixed-effects design to our data, we obtain similar estimates that are only slightly reduced in magnitude. Our estimates imply that being highly educated (i.e., having at least 13 years of schooling) is associated with 2.5–3.5 years of additional life expectancy at age 60 compared with being low-educated (i.e., having fewer that 10 years of schooling). This estimate is similar across gender and across time. Further, limiting the studied sample to MZ twins barely affects the estimates.

Taken together, our results suggest that the relationship between schooling and mortality is not primarily generated through the influence of shared genetic and/or environmental endowments affecting both schooling and mortality. When we test for the influence of confounding factors within twin pairs, we replicate earlier findings that birth weight predicts schooling differences within twin pairs. However, our twin fixed-effects estimates of the relationship between schooling and mortality are robust to controlling for birth weight differences. In addition, our results suggest that height, our proxy for cognitive ability, is a significant predictor of within-twin-pair differences in schooling but that including it in the regressions does not affect our main result of a significant relationship between schooling and mortality.

Theoretical Framework

As early as the 1970s, Michael Grossman formalized the idea that human capital may affect health, and thus longevity. He modeled education as a crucial component in individuals' health production function. In the model, education enhances the efficiency of health production, so that well-educated people are able to squeeze out a greater health output from a given health input (Grossman 1972). In a related hypothesis, Rosenzweig and Schultz (1982) suggested that educated people are instead able to choose a better mix of health inputs in the production of health.[2]

In essence, these theories are based on the premise that schooling in general better equips individuals to obtain, process, and value information—particularly on health-related behavior—and to translate this information into their lifestyle and habits, ultimately positively affecting their health and longevity (Kenkel 1991).

Not everybody was convinced, however, by the Grossman model's treatment of education as exogenously given. Fuchs (1982) argued that education and health may be related through unobserved additional factors, such as time preferences. People who are more future-oriented, ambitious, or disciplined ought to be more prone to forgoing instant pleasures and investing more in both education and health, given that both types of investments yield long-run benefits.[3] This would induce a positive correlation between education and longevity, without any causal relation between the two. Besides character traits reflecting time preferences, ambition, and discipline, genes and early-life environment may also be factors affecting both education and health,

[2] Economists have recently considered the relationship between education and a range of nonmonetary outcomes, such as health, criminal behavior, marriage, and political participation (see Lochner 2011 for an overview). This recent evidence points to the importance of education for a much wider range of outcomes than those observed on the labor market and thus suggests that the value of educational investments in society may be greater than previously thought.

[3] Later empirical research has found limited support for this specific mechanism (Cutler and Lleras-Muney 2010).

partly through influencing the formation of the aforementioned character traits. In fact, because a substantial part of the observed variation both in education and in mortality is believed to be genetic, it does not seem farfetched to assume that there could be some genetic overlap between the two.

Recent evidence also points to the importance of early-life conditions, including childhood health, for the production of education capital (Cunha and Heckman 2007). It is straightforward to conceptualize an education production function in which poor health early in life interferes with learning and is also associated with health later in life. Some evidence suggests, for instance, that low birth weight is associated with less schooling (Behrman and Rosenzweig 2004; Black et al. 2007; Oreopoulos et al. 2008; Royer 2009).[4] Health during childhood and adolescence has also been found to be important for later-life outcomes (Currie et al. 2010; Lundborg et al. 2011). Other mechanisms may be at work as well. Expected longevity may affect the incentives to invest in education, given that lower mortality means more years to reap the benefits of schooling (see, e.g., Jayachandran and Lleras-Muney 2009).

Health and mortality may also be associated with education if well-educated people are sorted into occupations with lower health risks, more fulfilling work tasks, and greater job and (by extension) life satisfaction. The link between education and mortality may also be mediated via higher social status or earnings expanding the consumption boundaries and possibilities for investing in health-promoting activities among well-educated people (see, e.g., Mirowsky and Ross 2003; Torssander and Erikson 2010).

To the extent that twins share important traits related to both schooling and mortality, such as genes, time preferences, and abilities, the twin fixed-effects design is helpful in estimating the causal effect of schooling regardless of whether the effect is a direct function of increased knowledge or mediated via sorting into varying kinds of occupations. If they differ on important traits, however, the estimates might be biased. In this article, we account for differences in birth weight and adult height. Comparing the twin estimates with and without such controls reveals something about the importance of early health differences and cognitive ability differences between twins.

Data

Our data originate from the Swedish twin registry, held by the Karolinska Insitutet in Stockholm. It is the world's largest twin registry and contains information on about 170,000 twins. We use data on the twins who were born in the period 1886–1958 and are older than 40 years, yielding a sample of about 50,000 same-sex twins, of which 18,000 are MZ.[5] The twin registry assigns zygosity based on questions about intrapair similarities in childhood. Several validations with DNA tests have shown this method to have 98 % or higher accuracy (Lichtenstein et al. 2002).

[4] Other twin studies, however, have found no significant association (Bonjour et al. 2003; Miller et al. 2005; Petersen et al. 2009).

[5] We impose the upper limit of 1958 because data for more recent cohorts of twins were not available to us.

The method used to collect the twin data gives rise to some selection issues, which are most important for the cohorts born in 1886–1925, given that only complete twin pairs who survived until the first survey in 1961 in these cohorts are included in our data. It seems likely that these "survivors" were, on average, healthier than the average twin born during the period. In the Empirical Model section, we discuss the potential bias that may arise from this selection. Under reasonable assumptions, we argue that our twin fixed-effects design largely deals with this selection problem and that any remaining selection bias gives rise to downwardly biased estimates.[6]

Because the data contain personal identifiers, we can link the data to various administrative registers. Key among these is the death register, which records the date and cause of death for all individuals who held a civil registration in Sweden and died in Sweden or abroad. Because we have information from the register for the 1961–2009 period and all twins in our sample were alive in 1961, we can observe all deaths that took place during this period for our twin sample.[7]

Our data on schooling come from several sources. For twins alive in 1990, we have access to register-based data on education from 1990 and 2007, from which we impute years of schooling.[8] In line with much of the previous literature, we focus on years of schooling in our main analysis, making our results easily comparable with those of other studies. In additional analyses, we provide results using educational categories.

Using register data on schooling should ensure that measurement errors are small. Some measurement error is likely to arise, however, because we impute years of schooling from highest obtained degree. Moreover, for individuals repeating or skipping grades, years of schooling will be inaccurately measured. Grade skipping or repetition was not very prominent in the Swedish school system during the study period, making this less of a concern. In 1946, for instance, only 1.6 % of all students repeated a grade (Ahlstrom et al. 1986).

[6] Selection issues are much less serious for those born in 1926–1958: average life expectancy had increased, and most twins would have survived until 1972–1973, when the first survey was conducted for these cohorts.

[7] The register does not cover Swedes who emigrated from Sweden and who no longer hold a civil registration in Sweden, however. Because we classify individuals who do not have a death date by 2009 in our data as still alive, we risk wrongly classifying individuals who emigrated and died abroad as still being alive. We believe this to be of little concern in our analyses, however, given that the number of emigrants in the cohorts used in our analyses is small. Although we cannot observe emigration directly in our data, we can draw on external sources. Statistics Sweden calculated the number of Swedes living abroad in 2003 for different age groups between 0 and 80 years (Nilsson 2004). For those aged 45 to 80 in 2003 (i.e., born between 1923 and 1958), the number of emigrants ranged from 250 among the early cohorts to 1,500 among the later cohorts; that is, only a small fraction of each birth cohort had emigrated. This fraction in our sample is substantially smaller because inclusion in our sample is conditional on surviving to 1972 and living in Sweden in 1972. We checked the number of potential emigrants by investigating how many individuals in our sample who lack income records in the last year of our income data (i.e., in 2007) had not died by 2007. The income records cover every source of income, including pensions. If there are no data, it is likely that the person either is dead or has emigrated. With this procedure, we found that 0.72 % (348 individuals) have no income and no death date. We reran our main regressions excluding this group, but the results were basically unchanged (results available on request).

[8] The information in these registers is reported from the various educational institutions directly to Statistics Sweden. Using the Swedish Standard Classification of Education (Svensk Utbildningsnomenklatur, SUN 2000), we impute years of schooling as follows: 6 for old primary school ("folkskola") (born before 1935), 7 for old primary school (born 1935 onward), 9 for (new) compulsory primary school, 10 for one year of high school, 11 for two years of high school, 12 for three years of high school, 13 for one year of university studies, 14 for two years of university studies, 15 for three years of university studies, 16 for four years of university studies, 17 for five years of university studies, 18 for a licentiate, and 20 for a PhD.

We can assess years of schooling from the register data for 71 % of our sample. For individuals with missing register data, we use information from surveys conducted by the Swedish twin register or from the 1970 census. For the cohorts born in 1926–1958, we use self-reported educational information from a survey conducted in 1973 as well as from the 1970 census. For cohorts born in 1886–1925, we use self-reported information from a survey conducted in 1963 as well as information from the 1970 census. Using the census data for the 1886–1958 birth cohort, we gain information on schooling for an additional 10 % of the sample.[9] By exploiting data from the surveys in 1973 and 1963, we obtain information for an additional 14 %. Thus, we have education data for 95 % of our sample.[10]

Birth weight information is taken from the birth records of twins born between 1926 and 1958. These data come from a part of the Swedish registry called BIRTH, which contains birth records of all twins born in Sweden between 1926 and 1958. Because of the sampling scheme, the birth weight information is available only for twins who survived until 1972. Still, we have access to about 25,000 observations on birth weight. Information on height, which we use as a proxy for cognitive ability, is available for the larger part of the sample. Questions about height were asked at multiple waves of the twin surveys.

Table 1 shows descriptive statistics for some of the key variables in our analyses. The mean number of years of schooling for MZ and DZ twins is 9.6 for males and 9.3 for females—a low average reflecting that a substantial proportion of the older cohorts have six years of schooling. Average years of schooling are slightly higher for MZ twins than for DZ twins.

The table also shows a relatively large number of deaths in our data. Among MZ and DZ twins, 41 % of males and 38 % of females died. For MZ twins only, the figures are slightly lower; 40 % among males and 35 % among females. The high fraction of deaths is a unique feature of the data that comes from having access to old cohorts. The mean death age for MZ and DZ twins with observed spells is 73.8 among males and 77.5 among females. Among the sample of MZ twins only, the mean death age is just slightly higher.

Because the twin fixed-effects design effectively exploits variation in schooling and mortality within twin pairs, it is crucial to have a large enough number of discordant twin pairs along these dimensions (see Boardman and Fletcher

[9] To impute years of schooling from the information in the 1970 census, we use the SUN codes in the census, which are somewhat less detailed than the SUN 2000 codes used in the 1990 and 2007 registers, together with information from Statistics Sweden. We assign years of schooling as follows: 6 for old primary school ("folkskola") (born before 1935), 7 for old primary school (born in 1935 and onward), 9 for (new) compulsory primary school, 11 or two years of high school, 12 for three years of high school, 14 for two or fewer years of university studies, 15 for three or more years of university studies, and 18 for a licentiate or a PhD.

[10] In the 1973 survey, education was coded in the same way as in the 1970 census, and we impute years of schooling in the same way. For the 1961 survey, we assign years of schooling for the majority of the sample based on the self-reports. We impute the following way: 6 for old primary school ("folkskola") (born before 1935), 7 for old primary school (born 1935 and onward), 9 for (new) compulsory primary school, 9 for "realskola," and 12 for secondary school. For three of the categories ("occupational school," "other," and "unknown"), we are unable to directly impute years of schooling. We calculate the average years of schooling associated with these three categories for the subsample of people for whom we have data on education from both the 1990 register and the survey in 1963. We then impute these averages for the individuals for whom the only source of education data is the 1963 survey. We use this imputation procedure for only 2 % of the sample.

Table 1 Descriptive statistics: Data from the Swedish twin registry

	MZ and DZ Twins		MZ Twins	
	Males	Females	Males	Females
Years of Schooling	9.589	9.306	9.744	9.592
	(3.320)	(3.186)	(3.357)	(3.239)
Mean Death Age (noncensored)	73.80	77.53	74.04	77.60
	(11.70)	(12.03)	(11.91)	(12.00)
Mean Censoring Age	65.21	67.41	65.07	66.68
	(9.924)	(11.07)	(9.759)	(10.81)
Fraction of Observed Deaths	0.590	0.621	0.604	0.650
	(0.492)	(0.485)	(0.489)	(0.477)
Number of observations	23,126	26,446	8,602	10,084
Birth Weight (g)	2,704.9	2,581.7	2,618.0	2,494.3
	(509.8)	(501.0)	(495.5)	(490.3)
Number of observations	12,365	13,007	4,676	5,266
Height (cm)	175.6	163.3	175.5	163.2
	(6.948)	(5.918)	(6.909)	(5.891)
Number of observations	19,596	23,449	7,462	9,030

Note: Standard deviations are shown in parentheses.

2015). We use information on death dates to measure mortality, and because no pair of twins has identical death dates, there is variation among all twins (except for the pairs in which no twin has yet died at the end of the study period). For education, our large sample comes in handy. In the sample of male MZ and DZ twins, 52 % of the twin pairs show a difference in schooling, commonly of one to four years. Among female MZ and DZ twins, the corresponding figure is 48 %. Although the fraction of pairs that differ in schooling in the MZ sample is somewhat smaller, a substantial fraction of pairs still differ.[11]

Cohorts vary substantially in the proportion with more than primary school, however. We observe a difference in schooling in 52 % of the twin pairs but in only 26 % of the twin pairs born in 1900. The corresponding number increases to 47 % for those born in 1925 and to 60 % for those born in 1958. The increasing variation reflects increased educational opportunities and implies less power for detecting a significant effect for the older cohorts because our twin fixed-effects estimator achieves identification from twin pairs that differ in schooling.

[11] Our figures on the fraction of pairs who differ in schooling corresponds relatively well with figures from previous studies using a twin design to study the returns to education (see, e.g., Ashenfelter and Rouse 1998; Holmlund et al. 2011).

A common concern when using twin data is external validity. For instance, twins are smaller at birth and may be different in other observed and unobserved ways. Table 2 displays comparisons of educational attainment between our twin population and the general population born during the same period. For the nontwin population, we use register data on entire cohorts of Swedes. We code education (in 1990) the same way in both populations. The comparison is complicated by the lack of surveys (other than the census) for the general population that can be used to assess educational levels for the older cohorts when register or census data are missing. Panel A of Table 2 shows comparisons for cohorts for which we can rely almost entirely on register data, thus making the comparison particularly "clean" (cohorts born in 1932–1958). Here, we see that the mean number of years of schooling is remarkably similar between the twin and nontwin population. There are some, albeit rather small, differences in the survival rates between the populations. For each follow-up period considered, twins are more likely to have survived, which likely reflects that inclusion in the twin sample is conditional on surviving until 1972. Panel B shows comparisons

Table 2 Descriptive statistics for twins and for the general population: Cohorts born in 1932–1958 and 1911–1931

	Twins		General Population	
	Males	Females	Males	Females
A. Cohorts Born in 1932–1958				
Years of schooling	10.91	10.92	10.87	10.87
	(3.02)	(2.85)	(2.99)	(2.83)
Deceased 1990	0.01	0.01	0.03	0.02
	(0.12)	(0.10)	(0.18)	(0.13)
Deceased 2000	0.04	0.03	0.07	0.04
	(0.20)	(0.17)	(0.25)	(0.20)
Deceased 2008	0.09	0.07	0.12	0.08
	(0.29)	(0.25)	(0.33)	(0.27)
Number of observations	13,178	14,122	1,776,163	1,667,472
B. Cohorts Born in 1911–1931				
Years of schooling	7.95	7.45	8.31	7.41
	(3.00)	(2.55)	(3.25)	(2.65)
Deceased 1980	0.06	0.04	0.07	0.04
	(0.24)	(0.19)	(0.26)	(0.19)
Deceased 1990	0.21	0.13	0.23	0.13
	(0.41)	(0.33)	(0.42)	(0.33)
Deceased 2000	0.46	0.31	0.47	0.32
	0.50	0.46	(0.50)	(0.47)
Number of observations	6,767	8,263	780,678	780,679

Note: Standard deviations are shown in parentheses.

for the older cohorts (those born in 1911–1931).[12] Here, the nontwin population of males appears somewhat more educated than the corresponding population of twin males. Among females, the differences appear tiny. Survival rates are very similar across the two populations. In sum, we do not find any evidence that the twin population differs greatly from the nontwin population born during the same period. In addition, if twins differ from singletons in unobserved ways, the twins fixed-effects methodology controls for those differences, just like they control for other endowments.

Empirical Model

Because the dependent variable measures the duration until death and there is prominent censoring among the younger cohorts, we use Cox models in the empirical analysis.[13] We also employ extended Cox models allowing for each twin pair to have a separate baseline hazard. Such models, sometimes referred to as stratified partial likelihood (SPL) models (which is a terminology we'll use henceforth), allow us to eliminate the influence of omitted variables operating at the twin pair level.[14] To see how, consider first the twin pair–specific baseline hazard:

$$\theta_{0p}\left(t, \eta_p\right), \tag{1}$$

where p denotes the pth twin pair, and η_p denotes unobserved heterogeneity at the twin pair level. Assuming proportional hazards, the hazard rate with twin pair–specific effects can then be written as follows:

$$\theta_p\left(t; x, \eta_p\right), (\beta) = \theta_0\left(t, \eta_p\right) \cdot \theta_1(x; \beta), \tag{2}$$

where x includes explanatory variables, such as years of schooling, and t denotes time. The baseline hazard, θ_0, is an arbitrary function of unobserved twin pair–specific heterogeneity and duration dependence, whereas θ_1 is a function of observed characteristics x. This approach thus applies the traditional partial likelihood approach to strata representing units, such as firms, schools, families, or (as in our case) twin pairs. In our case, the risk sets are replaced by twin pair–specific risk sets. Because all unobserved twin pair–specific effects are irrelevant in a comparison of individuals belonging to the same twin pair, they cancel out from the expression for

[12] We do not include cohorts born before 1911 because the 1970 census does not include educational information for cohorts born earlier than 1911—that is, for those who were 60 and older in 1970. For comparability, we therefore restrict the comparison to cohorts born in 1911–1931.

[13] An alternative would have been to run fixed-effects regressions on the probability of surviving until certain ages or to study 5- or 10-year survival. We prefer duration models in our main specification because they better exploit the variation in the data. For 10-year survival, for instance, two related twins who die within the same 10-year period but on different dates would not contribute to the identification of the estimates in a fixed-effects model. In our preferred duration model, however, any differences in the death dates contribute to the identification of the estimates. To shed more light on the magnitude of our estimates, however, we also exploit fixed-effects models.

[14] See Ridder and Tunali (1999) for a thorough exposition of the SPL model.

the likelihood. We therefore effectively rely on within-twin-pair variation in the covariates to identify the regression coefficients.

All observable covariates that do not vary within twin pairs, such as age and gender, also cancel out in the SPL model. In our main specification, we include only years of schooling as an explanatory variable. We do not include variables such as marital status, income, or occupation because we want to allow the effect of schooling on mortality to run through exactly such mechanisms. In our empirical analyses, we compare the results from a standard Cox model with those obtained from the Cox model with twin pair–specific baseline hazards.

The SPL estimator controls for the influence of unobserved variables at the twin pair level, much like the standard fixed-effects estimator. While the SPL estimator removes twin pair–specific unobserved heterogeneity, there may still exist environmentally induced unobserved differences within twin pairs that may relate to differences in both schooling and longevity. One threat to our particular application of the twin fixed-effects design would be the existence of unobserved health differences within twin pairs early in life. As discussed in the Theoretical Framework section, a number of twin studies have reported a significant effect of birth weight on schooling, although the evidence is mixed. A significant association between birth weight and schooling would be a problem if there were also a significant association between birth weight and mortality. Here, the evidence is very limited because such studies put great demands on the data, following people across the life cycle. We attempt to deal with this potential threat by running regressions on the subsample for which we have access to measures of birth weight and examine whether the inclusion of birth weight affects the relationship between schooling and mortality. If not, we can at least rule out that birth weight is responsible for the relationship between schooling and longevity.

Another potential threat to our twin fixed-effects design is unobserved ability differences within twin pairs. If such differences are related to both schooling and mortality, our estimates would most likely be upwardly biased because cognitive ability is positively related to both schooling and mortality. Very little evidence exists on the extent to which twins differ in ability and the extent to which such differences predict schooling differences. Sandewall et al. (2014), however, showed that IQ differences within twin pairs predict schooling differences.[15]

Several epidemiological studies have reported statistically significant correlations between early-life IQ and later-life health and mortality (see, e.g., the review by Batty et al. 2007). Holsinger et al. (2007) obtained the same result, but when relating differences in IQ between twins to differences in longevity between the same twins, they found that the estimated relationship became weak and insignificant. In addition, Lager et al. (2009) conducted a long-term follow up study of 1,500 children in which each additional year of schooling reduced the mortality hazard by 9 % to 12 % (Lager et al. 2009). When adding a measure of childhood IQ, however, they found that the relationship between schooling and mortality was not affected.

Although we cannot observe cognitive ability or IQ in our data set, we can control for variables that are known to be correlated with IQ. One such variable is

[15] One cannot rule out that IQ differences between twins are generated through earlier differences in schooling inputs, such as teachers. In fact, this is what two Swedish studies found when using the same data Sandewall et al. (2014) used (Carlsson et al. 2013; Meghir et al. 2013).

height, which has been shown to be overall strongly related to childhood IQ (Case and Paxson 2008; Lundborg et al. 2014; Persico et al. 2004), although less so when differencing among twins (see, e.g., Beauchamp et al. 2011; Keller et al. 2013). Height is often taken as an indicator of early-life health and can thus be expected to reflect variation in both childhood health and IQ. If the results change drastically when controlling for this proxy, however, we would be concerned that unobserved differences in cognitive ability are influencing the results. We fully acknowledge that this by no means is a perfect solution. Because height is also related to early-life health, any change in the results when accounting for height could obviously be interpreted as picking up health differences. However, an absence of change in the result could provide some indication that differences in cognitive abilities play no great role, although we then need to assume that the variation in IQ that is relevant for mortality is picked up by our height measure.

Twins with different levels of schooling may differ in other, unobserved ways. Our estimate of schooling would still indicate a causal effect, however, if these unobserved factors are not the same as those that determine longevity. In our twin fixed-effects context, we cannot test for this possibility.[16] However, we believe we can at least rule out some of the common suspects, such as birth weight and (via height) cognitive ability to some extent. To shed light on other possible confounders, we focus on twin pairs who differ little in their levels of schooling, among whom differences in unobserved factors should also be fewer. If the estimates do not change much when we rely only on twin pairs that do not differ much in schooling, this would then suggest that unobserved factors play less of a role. In the same spirit, we can run regressions that include only twins who are very similar in terms of height and birth weight.

Another issue that arises when interpreting twin-based estimates is that parents may reinforce or compensate for differences in endowments between twins. In the former case, the importance of such differences would be magnified. Most studies on data from Western countries seem to support the hypothesis that parents try to compensate for ability differences between children (for an overview of this literature, see Almond and Currie 2011).[17] To the extent that cognitive ability is associated with underlying health—or to the extent that parents also compensate by investing more in the education of their less-healthy offspring—our estimates of the impact of schooling on mortality will be downwardly biased. Without data on parental behavior, it is obviously impossible to address these concerns beyond trying to control for observable differences that may induce or reflect parental responses, such as birth weight and height.

As mentioned in the Data section, the sampling of the twins may introduce some selection bias. For twins born in 1886–1925, only those alive in 1961 are included in the sample. For the older cohorts, we are thus likely to observe a sample of unusually healthy individuals. Moreover, it is plausible that that a low-educated individual who is still able to survive to 1961 has better unobserved underlying health, on average, than a highly educated individual from the same birth cohort. In this case, a negative correlation between schooling and certain unobserved traits, such as healthy genes and early-

[16] In principle, this possibility is testable if a good instrument exists to generate random variation in schooling between members of twin pairs. To the best of our knowledge, there are no twin data containing such instruments.

[17] In contrast, a few studies found that parents reinforce endowment differences between children (see, e.g., Behrman et al. 1994).

life health, would arise, leading to a downward bias in the estimate of schooling. Here, the twin fixed-effects design gives us an advantage: many such unobserved traits are shared by twins in a given twin pair and therefore cancel out, which likely reduces the importance of the selection problem. We will shed light on this phenomenon in a sensitivity analysis.

Certain selection issues may still prevail, however. For instance, if a lower-educated twin is likely to have better underlying health than the higher-educated co-twin in pairs where both survive until an older age, this would push the estimates downward. Moreover, because survivors can be expected to be healthier, on average, the effect of schooling may be smaller for this group if there is decreasing marginal productivity in the production of longevity with respect to schooling. However, these selection issues are much less important for the younger cohorts. Therefore, in the sensitivity analysis, we examine the extent to which the results differ by cohort and by age.

The estimated association between schooling and mortality would also be downwardly biased if there are spillover mechanisms of knowledge between twins, whereby the lower-educated twin may learn from and imitate the behavior of the co-twin, who in turn has gained knowledge via more years of schooling (e.g., Boardman & Fletcher 2015; Kohler et al. 2011).[18]

Finally, one needs to assume that investments in schooling do not depend on the future moment of death of the main person, given x_i and η_p. This assumption would fail, for instance, if the main respondent early in life has some idea about his or her expected length of life. If this expectation is low, he or she may then invest less in schooling, and what looks like an effect of schooling on mortality may then just reflect reverse causality. We expect such anticipation effects to be weak, however, because a useful feature of the twin fixed-effects design is that any signaling value that parental health or longevity may have for the twins' own expected longevity is shared by the twins. Differences in early-life health could lead to differences in expectations about longevity, however. Here, we expect the regressions including controls for height and birth weight to be informative: variation in these factors could generate differences in expectations about longevity. To preclude the influence of more severe negative early-life health shocks, we limit the sample to those twin pairs in which both survived to at least age 40.[19]

Results

We begin by analyzing the relationship between years of schooling and mortality. Table 3 shows the Cox regression results for pooled MZ and DZ twins and for MZ twins only, with and without allowing for twin pair–specific baseline hazards. In the models without twin pair–specific baseline hazards, we adjust for gender and birth year. Contrasting the twin fixed-effects results obtained for the pooled sample of MZ and DZ twins with those obtained for MZ twins only is of interest because it might reveal

[18] The opposite would be true, implying upward bias, if the twin with less schooling rebelled by behaving differently from the better-educated co-twin.

[19] For comprehensive overviews of common criticisms to twin studies and the circumstances under which the estimated results are unbiased, see Amin et al. (2015b), Boardman and Fletcher (2015), and Kohler et al. (2011).

something about the role of genetic endowments for the relationship between schooling and mortality. In the pooled MZ and DZ sample, not all genetic endowments will be washed out; by contrast, in the MZ estimates, within-twins differences in genetic endowments will be completely washed out. Thus, if these endowments are important confounders, the MZ estimates should differ from the combined MZ/DZ estimates.

As shown in the first column of Table 3, the hazard ratio of 0.960 is significant and implies that one additional year of schooling is associated with about a 4 % lower mortality hazard in our pooled sample of MZ and DZ twins. This estimate is substantial but somewhat lower than the estimates obtained in some previous studies using Swedish data (e.g., Lager et al. 2009).

The second column displays the results of allowing for twin pair–specific baseline hazards in our sample of MZ and DZ twins. Here, any unobserved factors that may affect schooling and mortality that are shared by a pair of twins are cancelled out. The hazard ratio slightly decreases to 0.954, implying a somewhat greater effect. Thus, unobserved common early-life environment and shared genes do not appear to greatly bias the estimates obtained in column 1. Admittedly, DZ twins, who constitute the major part of the pooled sample, share only 50 % of their genes, but the fact that the estimates hardly change when moving to the SPL model speaks against any strong influence of genetics in the relationship between schooling and mortality. This conclusion is supported when we restrict the sample to MZ twins only, as shown in columns 3 and 4. Here, the results are indeed very similar to the estimates obtained for the pooled sample of MZ and DZ twins.[20,21]

Since the Cox model assumes proportional hazards, we test for this using the standard test that examines the scaled Schoenfeld residuals (Grambsch and Therneau 1994). The test rejects the proportional hazards assumption in all the models in Table 3. Given that the hazards are not proportional, the estimated parameters still yield a weighted average of the considered effect over the studied time frame, but they will conceal how the effect changes with time (or, in this case, age). A particularly problematic scenario is if the hazards cross, implying that a characteristic that is positively associated with the outcome in the beginning of the time frame becomes negatively associated with it over time. In order to check which interpretation is correct, we classified individuals in into high-educated (more than 12 years of schooling) and low-educated (12 or fewer years of schooling) and visualize the hazards in Fig. 1 (adjusted for gender). There is no evidence of the hazards crossing, but the hazards tend to converge at young and old ages. Hence, the results so far should be interpreted as giving average effects over the studied age span, rather than as measuring an effect that is constant with age. Later, we investigate potential age differences in more detail by dividing the sample into age groups; as we show there, the proportional hazards assumption then holds up.

What could explain the small difference in results when accounting for all possible twin pair–specific factors? One can argue that the Cox estimates and the SPL estimates are not strictly comparable because SPL estimates are based only on twin pairs who

[20] We tested for statistically significant differences in the coefficient of schooling between MZ and DZ twins by pooling the samples and including an interaction term between schooling and being a MZ twin. This interaction term was insignificant, however (p = .34).

[21] The similarity in results with and without including twin fixed effects does not mean that genes are not important determinants of schooling or mortality. Instead, it suggests that the genes affecting schooling are not the same as those affecting mortality.

Table 3 Cox regression on mortality and education: Data from the Swedish twin registry

	MZ and DZ Twins		MZ Twins	
	Cox	SPL	Cox	SPL
A. Pooled Sample				
Years of schooling	0.960**	0.954**	0.958**	0.955**
	(0.954–0.965)	(0.941–0.968)	(0.949–0.967)	(0.929–0.981)
Number of observations	49,572	49,572	18,686	18,686
B. Males				
Years of schooling	0.960**	0.957**	0.963**	0.948**
	(0.953–0.967)	(0.939–0.975)	(0.951–0.975)	(0.914–0.984)
Number of observations	23,126	23,126	8,602	8,602
C. Females				
Years of schooling	0.959**	0.951**	0.951**	0.962†
	(0.951–0.967)	(0.930–0.972)	(0.937–0.965)	(0.925–1.001)
Number of observations	26,446	26,446	10,084	10,084

Notes: Confidence intervals are shown in parentheses. Models 1 and 3 are standard Cox models. Models 2 and 4 are Cox stratified partial likelihood (SPL) models. Models in columns 1–3 include controls for gender and birth year (dummy variables).

$^{\dagger}p < .10$; $^{**}p < .01$

differ in schooling, whereas the Cox estimates use variation between and within twin pairs. Clearly, if the twin pairs that differ in schooling are systematically different from other twin pairs, the estimates that are based on them may not be generalizable to the full twin population and they have a local average treatment effect (LATE) flavor.[22] We check this by also running Cox regressions only on twin pairs that differed in schooling. This generates a hazard ratio of 0.964, which is very similar to the estimate of 0.960 obtained for the Cox model on the full population.[23] We obtain a similar pattern when focusing on MZ twins only; here, the Cox estimates based on only the twins who differ in schooling generates a hazard ratio of 0.960. Thus, the difference in magnitude of the estimates between the Cox and SPL models remains small, and the main message of the analysis does not change; a statistically significant relationship between schooling and mortality also remains when we account for shared genes and environment.

Before proceeding with gender-specific results, we want to make sure that our main results are not driven by some outliers in the data. As Amin (2011) showed, twin fixed-effects estimates are especially sensitive to the influence of a few outliers in the

[22] However, an often mentioned potential advantage of within-twins estimates is that they might have less of a LATE flavor than IV estimates that depend on changing minimum schooling or school-leaving age requirements. The reason is that the twin fixed-effects estimates use variation in schooling over the entire schooling distribution, whereas many IV studies only use limited variation in schooling induced by the instrument (Amin et al. 2015b; Behrman et al. 2011).

[23] We also checked whether the twin pairs that differed in schooling also differed from other twin pairs in terms of observable characteristics. The results suggested that differences within twin pairs have become more common over time, as the average birth year of twin pairs that differed in schooling was 1937 compared with 1927 for twin pairs who did not differ. This difference was statistically significant. We return to the issue of changes in the estimates over cohorts/ages in subsequent sections.

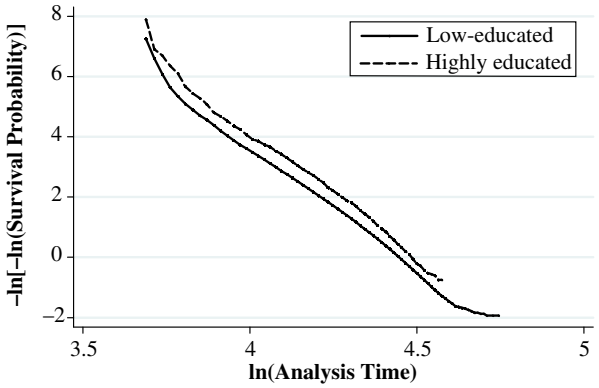

Fig. 1 Illustration of the proportional hazards assumption

explanatory variable. We therefore rerun our results first excluding all twin pairs in which the twins differed by more than four years in terms of schooling. Doing so actually decreases the hazard ratio somewhat, to 0.94. Further restricting the sample to pairs with a schooling difference of no more than three years yields similar results (hazard ratio of 0.92), as does restricting the sample to those with at most a two-year difference. These results are also of interest for another reason: it appears reasonable that twin pairs who differ less in terms of schooling also differ less in unobserved factors that could potentially bias our results. If these unobserved factors were of great importance, however, we would expect the estimates based on twin pairs with similar schooling to be of lower magnitude. This is not what we find, however.[24]

Results by Gender

We next consider the possibility that the estimates differ between males and females by running separate regressions by gender. Given the different roles of males and females in society, it would not be surprising to find differences in the nature of the relationship between mortality and schooling.

Panel B of Table 3 shows the results for males. Allowing for twin pair–specific baseline hazards via SPL estimation in the sample of male twins, we now obtain a hazard ratio of 0.957, whereas the corresponding Cox estimate is 0.960. When we restrict the sample to MZ twins only, the SPL estimate decreases somewhat (0.948).

For females, the SPL hazard ratio of 0.951 in the pooled sample of female twins is close to that of the male twins, and the difference is not statistically significant.[25] When we restrict the sample to MZ twins only and apply the SPL estimator, the hazard ratio increases somewhat to 0.962. The somewhat greater effect of schooling obtained for males in the SPL model is in line with previous evidence of a somewhat stronger gradient among males, although our results suggest that the differences are quite small.[26]

[24] We also ran regressions excluding people who survived the ages of 100 and 90, respectively. We lost about 3,000 individuals by imposing this restriction, but the results barely changed.

[25] When we pool genders and interact an indicator for being male with schooling, the interaction term is insignificant in the SPL model ($p = .68$).

[26] The differences are not statistically significant, as revealed by insignificant interaction terms in both the SPL and in the standard Cox model ($p = .68$ and $p = .17$, respectively).

Internal Validity of the Twin Fixed-Effects Design

Having established a set of baseline results, we next turn to the interpretations of our data and results. We focus on the internal validity of our twin estimates by conducting a series of tests for potential confounding factors. In particular, we are concerned that schooling achievement within twin pairs could be correlated with factors that also determine longevity and that are unobserved to the analyst, thus not warranting a causal interpretation of our estimates.

Birth Weight, Schooling, and Mortality

First, we examine the extent to which controlling for birth weight, our proxy for early-life health, affects our results. Because birth weight information is collected only for those born in 1926–1958, the sample size is now halved, from 50,000 to about 25,000.

We start by checking for any evidence that birth weight actually affects schooling in our sample. This serves as a credibility check of our birth weight measure, given that previous studies have found significant effects of birth weight. As shown in the first column of Table 4, the relationship between birth weight and years of schooling is not significant in an OLS model, with controls for gender and birth cohort. When we add fixed effects to our pooled sample of MZ and DZ twins, however, the results show that a 100 g increase in birth weight is associated with 0.023 more years of schooling.[27] This estimate corresponds closely to Royer's (2009) estimate, obtained using U.S. data, that a 100 g increase in birth weight is associated with 0.02 additional years of schooling. When we restrict our sample to MZ twins, the effect actually increases in magnitude. To compare the results with those of Black et al. (2007), we also estimate the effect of log birth weight on high school completion. The estimates show that a 10 % increase in birth weight is associated with a 0.84 percentage point increase in the probability of high school completion, which is very similar to the 0.9 percentage point increase Black et al. (2007) obtained. The similarity of our findings to those obtained in previous studies suggests that our birth weight measure is valid.

Our finding of a positive association between birth weight and schooling raises some concerns about our main results. If birth weight is also positively related to longevity, then our estimates based on the full sample, for which we lack birth weight information for all, might be upwardly biased. We therefore reestimate our main mortality models on the sample with birth weight information to check how the estimate of schooling is affected by controlling for birth weight. We first establish whether schooling and mortality are also significantly associated in this restricted sample with information on birth weight. Indeed, this is the case for our pooled sample of MZ and DZ twins: as shown in the first and third columns of Table 5, the results are similar to our baseline results. For MZ twins, the estimate is similar in magnitude but does not reach statistical significance in this smaller sample.

[27] We also tried alternative specifications of the birth weight variable, such as an indicator of low birth weight (birth weight <2,500 g). The indicator of low birth weight is significant, but when we included both this indicator and the birth weight variable, the latter always dominated and the former became insignificant.

Table 4 Regressions on birth weight and education (panel A) and height and education (panel B): Data from the Swedish twin registry

	No Fixed Effects		Fixed Effects	
	All Twins (1)	MZ Twins (2)	All Twins (3)	MZ Twins (4)
A. Birth Weight (100 g)	0.002	0.013**	0.023**	0.042**
	(0.004)	(0.006)	(0.006)	(0.008)
Number of observations	25,372	9,942	25,372	9,942
B. Height	0.058**	0.063**	0.019**	0.022**
	(0.002)	(0.004)	(0.003)	(0.007)
Number of observations	43,042	16,492	43,042	16,492

Notes: Standard errors are shown in parentheses. Columns 1 and 2 report OLS regression results with controls for gender and birth year. Columns 3 and 4 report linear fixed-effects regressions.

**$p < .01$

As shown in columns 2 and 4 of Table 5 (panel A), the magnitude of the hazard ratio for schooling is almost unchanged when we account for birth weight differences. We also again tried other indicators of birth weight, such as an indicator of low birth weight (less than 2,500 g) and log birth weight, but the results did not change. Our results thus suggest that although birth weight predicts schooling, the omission of birth weight in our main regressions is unlikely to bias the estimate of schooling to any important extent.

Table 5 Regression on mortality and years of schooling, controlling for birth weight (panel A) and controlling for height (panel B): Data from the Swedish twin registry

	MZ and DZ Twins		MZ Twins	
A. Birth Weight Sample				
Years of schooling	0.936**	0.936**	0.961	0.960
	(0.013)	(0.013)	(0.025)	(0.025)
Birth weight (100 g)		0.978**		1.005
		(0.009)		(0.017)
Number of observations	25,372	25,372	9,942	9,942
Control for birth weight	No	Yes	No	Yes
B. Height Sample				
Years of schooling	0.957**	0.957**	0.956**	0.956**
	(0.008)	(0.008)	(0.014)	(0.014)
Height (cm)		0.995		0.994
		(0.003)		(0.008)
Number of observations	43,042	43,042	16,492	16,492
Control for height	No	Yes	No	Yes

**$p < .01$

Proxies for Cognitive Ability Differences

We next turn to the question of potential differences in cognitive ability between twins. As discussed in the Empirical Model height is related to childhood and adolescent IQ as well as to early-life health and nutritional inputs. Thus, an absence of a change in results when we control for a variable that is known to be strongly correlated with childhood IQ would be informative.

First, we establish that height and schooling are significantly related in our sample. As shown in panel B of Table 4, a 10 cm increase in height is associated with a 0.06 increase in the number of years of schooling. Because height information is present for the lion's share of the full sample, we are no longer restricted to the smaller sample size used for assessing the role of birth weight. When we impose twin pair–specific baseline hazards, the estimate decreases to about one-third (0.019) but remains significant at the 1 % level, as shown in the second column. Similar and significant estimates are obtained if we restrict the sample to MZ twins.

To check whether our main result on the relationship between schooling and mortality changes when we control for height in the regression, we first estimate the relationship between schooling and mortality in the part of the sample for which we have information on height. As seen in panel B of Table 5, the resulting estimates are very similar to those obtained for the full sample. When we also include height in the regression, the results hardly change. Moreover, the results show no significant relationship between height and mortality in either model, which is an interesting result in its own right.

Because height is known to be correlated with cognitive ability, an alternative sensitivity check would be to examine whether the results change when we include only twin pairs who are very similar in terms of height and who, thus, should also be more similar in terms of cognitive ability and early-life health. We rerun our SPL regressions on a sample of 20,104 MZ and DZ twin pairs who differ by no more than 2 cm in height. The estimated hazard ratio is still significant and of a magnitude similar to that for the full sample (0.962). The estimate is also similar, and significant, if we restrict the sample to twins who differed by no more than 1 cm (0.969, n = 13,334).

Learning From DZ Twins

We have shown that controlling for birth weight differences and differences in height, which arguably picks up some variation in cognitive ability, does not affect our results to any great extent. Of course, the possibility remains that other unobserved factors that differ between twins will bias our estimates. We can shed some light on this by exploiting the information on zygosity in the data. It is plausible that MZ twins are more alike than DZ twins not only in terms of schooling but also in terms of preferences and personality traits. This is, for instance, implied in variance decomposition studies that try to single out the genetic component in such traits using contrasts between MZ and DZ twins (e.g., Cesarini et al. 2009; Jang et al. 1996). In our case, preferences that affect both schooling and longevity should be more alike among MZ twins than among DZ twins. If so, we would also expect the MZ estimates to be smaller than the DZ estimates, given that

a larger part of the influences of such preferences will be differenced out among MZ twins.[28] On the other hand, little change in the results would suggest that differences in unobserved factors, such as preferences, are less important.

When we explicitly make the comparison between DZ and MZ estimates, the hazard ratios barely differ (0.954 and 0.955). The same is true when we run separate regressions by gender. We can therefore rule out that unobserved genetic endowments play an important role in the relationship between schooling and mortality. Moreover, the result suggests that measurement errors do not play a great role because one would otherwise expect smaller MZ estimates.

Summing Up Internal Validity Concerns

We acknowledge that we are unable to completely rule out that the within-twin-pair relationship between schooling and longevity reflects the influence of omitted variables affecting both these outcomes. Our tests show, however, that some of the factors that determine schooling differences between twins—and that are often difficult to observe—do not seem to bias our estimates. We therefore believe that our results are consistent with a causal interpretation that schooling improves longevity. When we control for the twins' genetics and childhood environment, which are plausible candidates for determining both schooling and longevity, our estimates barely change. When we also control for some observable factors within twin pairs, such as early-life health and cognitive ability, our estimates remain unaffected.

Nonlinear Returns

Next, we consider the possibility of important nonlinearities in the relationship between schooling and mortality. For this purpose, we define three groups: those with less than 10 years of schooling, those with 10–12 years of schooling, and those with at least 13 years of schooling. The latter category includes people with a university exam.

The results in panel A of Table 6 show that MZ and DZ twins with 10–12 years of schooling have about 14 % lower mortality hazard than those with less than 10 years of schooling. This estimate is only mildly affected when we allow for twin pair–specific baseline hazards. For twins with a university exam or some university studies (at least 13 years of schooling), the estimates are more dramatic: the mortality hazard is more than 40 % lower in the SPL model (38 % lower hazard among MZ twins only).

Table 6 also shows estimates separately for male and female MZ twins. For the combined sample of MZ and DZ twins, the estimates for both men and women are very similar to those obtained for the pooled sample. Turning to MZ twins, it seems as if the

[28] This would be true as long as at least some part of the variation in such preference traits is due to genetics. This is also what the studies by Cesarini et al. (2009) and Jang et al. (1996) suggest. A difference in results between MZ and DZ twins could also be explained by the greater role that measurement errors play among MZ twins than DZ twins. However, the bias resulting from relying on purely self-reported measures for education in estimating the association between schooling and health has been found to be of limited importance (e.g., Amin et al. 2015a, b). In addition, because data on education in this study are taken from registers for a large part of our sample, we do not believe that measurement errors play a great role.

Table 6 Cox regression on mortality and education, nonlinear returns: Data from the Swedish twin registry

	MZ and DZ Twins		MZ Twins	
	Pooled	SPL	Pooled	SPL
A. Males and Females				
10–12 years of schooling	0.855**	0.849**	0.828**	0.814**
	(0.825–0.886)	(0.786–0.917)	(0.780–0.879)	(0.709–0.936)
13 or more years of schooling	0.634**	0.569**	0.660**	0.617**
	(0.591–0.680)	(0.487–0.665)	(0.590–0.739)	(0.463–0.823)
Number of observations	49,572	49,572	18,686	18,686
B. Males				
10–12 years of schooling	0.864**	0.870*	0.844**	0.838†
	(0.823–0.908)	(0.782–0.968)	(0.777–0.916)	(0.689–1.021)
13 or more years of schooling	0.635**	0.579**	0.689**	0.594**
	(0.578–0.696)	(0.473–0.710)	(0.592–0.801)	(0.404–0.875)
Number of observations	23,126	23,126	8,602	8,602
C. Females				
10–12 years of schooling	0.838**	0.827**	0.801**	0.789*
	(0.794–0.883)	(0.740–0.925)	(0.734–0.875)	(0.648–0.961)
13 or more years of schooling	0.625**	0.558**	0.635**	0.656†
	(0.561–0.697)	(0.438–0.712)	(0.534–0.754)	(0.427–1.010)
Number of observations	26,446	26,446	10,084	10,084

Notes: Confidence intervals are shown in parentheses. Models 1 and 3 control for sex and birth cohort. Models 3 and 4 are Cox SPL models.

$^{\dagger}p < .10$; $*p < .05$; $**p < .01$

effect of 10–12 years of schooling compared with less schooling is somewhat stronger among women.[29] All in all, the estimates look rather linear in the categories considered.

Selection and Changes in the Estimates Over Time

As argued in the Data section, the sampling used in the twin registry may induce some selection problems, which most likely leads to conservative estimates using the Cox model. If selection is an important issue, we would expect that the downward bias in the estimates is larger for the older cohorts, whose inclusion in the sample depended on a longer survival. A person born 1890, for instance, had to survive to the age of 71 to be included. We can therefore check the sensitivity of our results to the inclusion of the older cohorts. However, when we exclude older cohorts, we not only exclude cohorts for which more selection has taken place but we also are less likely to observe completed lifespans in the remaining cohorts. We will need to assume, for the moment, that the proportional hazards assumption holds—that is, that the effect of schooling remains constant over time (across age). Later, we will investigate this assumption as well.

[29] We also repeated the sensitivity tests of the previous sections, controlling for birth weight and height. The results were again robust.

Table 7 shows both Cox and SPL estimates with different restrictions on what cohorts to include. In the first row, we replicate the main results for ease of comparison. Excluding the oldest cohort, born in 1886–1895, does not cause any radical change in the results, as shown in the second row. Nor do dramatic changes occur when we exclude the cohorts born in 1896–1905 and 1906–1915. If selection bias had been important, we would have expected that the Cox estimates would grow in magnitude as the older cohorts were excluded. Moreover, we would expect the SPL estimates to differ more from the Cox estimates in the older cohorts, given that the SPL model alleviates some of the selection problem. This is not what we find, however; in fact, none of the coefficients reported so far differ in a statistically significant manner.[30]

We next turn to the results for the later born cohorts, where the results suggest that there have been some changes over time in the estimates. Including only cohorts born in 1936–1958 gives a stronger impact of schooling, with a hazard ratio of 0.91 in the SPL model. This pattern becomes even more pronounced when we restrict the sample to the even younger cohorts born in 1946–1958, where the hazard ratio is now 0.84. These results are consistent with many previous studies that have found that the socioeconomic gradient in mortality has grown stronger over time. For instance, well-educated people have been better able to take up new medical interventions (Glied and Lleras-Muney 2008). It is only for the latest period considered, 1946–1958, that we obtain a coefficient that is significantly different from the rest.[31]

The changes in estimates in the recent cohorts can also be explained in other ways. As mentioned earlier, the age composition of the sample automatically changes when we impose different cohort restrictions. For the younger cohorts, many more deaths are censored simply because fewer people have reached high ages. Thus, the cohort effect found in Table 7 could reflect that the impact of schooling varies by age. We can shed some light on the plausibility of this explanation by comparing the estimates for similar age groups who belong to different cohorts—that is, by moving the age window across time.

For this purpose, we focus on specific cohort subdivisions of the main sample for which we are able to make similar age groupings. If the changing age composition of the sample explains the changes in the results, we would expect to see no great differences when comparing the results for the same age groups across cohorts. On the other hand, a change in results across the same age groups across cohorts would suggest that real changes in the relationship between schooling and mortality have taken place. This again assumes that our SPL estimator eliminates selection issues and that any remaining ability bias remains more or less constant over time.

[30] We cannot rule out that a strong downward bias from selection is offset by a strong upward bias from unobserved ability in the older cohorts. In later cohorts, a weaker selection bias would then also be offset by a weaker ability bias. It appears likely to us that improved access to education over time leads to less rationing of education according to skills. Moreover, when credit constraints are less binding, parents could more easily afford to send both twins to higher education institutions, whereas more binding constraints may lead parents to invest more in the more able twin. We can shed some light on this possibility by examining if the effect of height, which partly picks up cognitive ability, on education decreases over time. This is not the case, however; if anything, the effect is greater for the post-1940 birth cohorts.

[31] We tested this by including interaction terms between schooling and birth cohort indicators in the SPL model. For cohorts born in 1946–1958, the interaction term was significant for males and females combined ($p < .01$) and for males ($p < .01$), but was not significant for females ($p = .36$).

Table 7 Cox and SPL regressions on mortality and with different sample restrictions: Data from the Swedish twin registry

	Pooled		Males		Females	
Year-of-Schooling Sample Restriction	Cox	SPL	Cox	SPL	Cox	SPL
1886–1958	0.960**	0.954**	0.962**	0.957**	0.955**	0.951**
n	49,572	49,572	23,126	23,126	26,446	26,446
1896–1958	0.959**	0.956**	0.962**	0.960**	0.955**	0.950**
n	48,346	48,346	22,630	22,630	25,716	25,716
1906–1958	0.958**	0.953**	0.961**	0.956**	0.954**	0.947**
n	45,406	45,406	21,350	21,350	24,056	24,056
1916–1958	0.955**	0.946**	0.959**	0.947**	0.950**	0.944**
n	39,572	39,572	18,692	18,692	20,880	20,880
1926–1958	0.947**	0.943**	0.953**	0.952**	0.937**	0.930**
n	31,686	31,686	15,174	15,174	16,512	16,512
1936–1958	0.923**	0.913**	0.924**	0.931**	0.922**	0.883**
n	24,388	24,388	11,818	11,818	12,570	12,570
1946–1958	0.878**	0.837**	0.867**	0.831**	0.896**	0.845**
n	14,306	14,306	7,038	7,038	7,268	7,268

Note: Standard errors are shown in parentheses.

**$p < .01$

Table 8 presents the SPL results from this exercise. The first two columns give the mortality hazard ratios for ages 50–60 in the cohorts born in 1922–1934 and 1935–1947 respectively; the subsequent columns give the corresponding hazards for later ages and comparable cohorts. The relationship between schooling and mortality within a particular age group is not radically different across cohorts. In the age group 50–60, however, the hazard ratio for the 1922–1934 and 1935–1947 cohorts is 0.93 and 0.90, respectively, which suggests that the effect within the same age group becomes somewhat weaker across cohorts. The difference in hazard ratios is not statistically significant, however, when tested with an interaction term ($p = .49$). For the other age ranges considered, estimates are even more similar across cohorts.[32] Thus, the estimates suggest that the relationship between schooling and mortality looks similar across cohorts but that it varies across the life cycle. The difference in results when we exclude cohorts (shown in Table 7) therefore reflects a weakening of the relationship between schooling and health by age rather than changes in the relationship across cohorts. The reduced importance of schooling by age could be explained by a mechanism whereby the underlying health of those surviving to higher ages is greater and schooling matters less for health production at higher levels of underlying health.

[32] The interaction terms between schooling and the cohort indicators for the age groups 60–70 and 70–80 were insignificant ($p = .63$ and $p = .91$, respectively)

The comparisons across age groups also shed light on the proportional hazards assumption. We showed earlier that the assumption does not hold when all age groups are pooled and that this likely reflected the convergence of hazards at higher ages. The results in Table 8 as well as our test of the proportional hazards assumption by age group confirm these findings.[33]

How Much Life Expectancy Is Gained by Higher Schooling?

The Cox and SPL survival models analyzed in this article are semiparametric in the sense that they analyze the impact of explanatory factors (e.g., schooling) but not the parameters of the underlying baseline hazard parameters. These models use the ordering of relevant events (deaths) but not information on how far apart or exactly when on the time scale different events occur. Hence, these models yield estimates of the mortality hazard ratios, but in the absence of a specified baseline, such hazards are not directly translatable in terms of more detailed longevity prospects. To capture such prospects via survival analysis, the baseline has to be specified and its parameters must be estimated.

Unfortunately, it is infeasible to use our SPL twin fixed-effects design to estimate a parametric proportional hazards model because there is simply not enough information (or degrees of freedom) from two observations to allow for specification of one specific baseline for each and every stratum or twin pair. Nevertheless, the results presented so far strongly suggest that unobserved characteristics operating at the twin level do not influence the estimated impact of schooling on mortality in any important way. To illustrate how the hazard ratios for the effect of schooling on mortality may translate into survival propensities, we therefore present estimates of mean remaining life expectancy at age 60, based on the Gompertz distribution.[34,35] These estimates are likely slight underestimates because our nonlinear SPL estimates shown are somewhat greater than the nonlinear Cox estimates.

The sample included in this analysis is restricted to only twin pairs who were both alive at age 60. Hereafter, we estimate Gompertz survival models for different subpopulations according to gender, educational group (less than 10 years, 10–12 years, or 13 or more years), and birth cohort (1900–1919 and 1920–1929), with birth year as a continuous explanatory variable.[36]

[33] We test the proportional hazards assumption using Schoenfeld residuals (see the *phtest* option in Stata). For the age groups 50–60, 60–70, and 70–80, the Prob > χ^2 are 0.55, 0.93, and 0.51, respectively.

[34] The Gompertz distribution has been shown to describe old age mortality rather well. This model is a parametric proportional hazard model in which the baseline hazard takes the following form:

$$\theta_0(t, a, b) = ae^{t/b}, \tag{3}$$

where a and e are shape and scale parameters, respectively, of the baseline; and t is, just as in the Cox and SPL models, the randomly distributed death time. Hence, the model can be expressed in terms similar to the Cox model with the addition that the baseline is specified according to a Gompertz distribution with specified shape and scale parameters.

[35] The main purpose of this exercise is not to exactly capture specific remaining life expectancies for different subgroups but rather to illustrate the differences in these expectancies—that is, to illustrate how the impact of education on mortality translates to longevity.

[36] On a computational level, the studied time/age profile is rescaled so that entrance in the relevant risk set occurs at time (t) 0 (at age 60) in order to more accurately capture old age mortality. The resulting estimates of the shape and scale parameters are then used to calculate life expectancies by integrating the random variable t with respect to its probability measure given by the estimated Gompertz density, over its support (0, inf)—that is, from age 60 ($t = 0$) onward.

Table 8 Cox SPL regressions on mortality and education for MZ and DZ twins by age group and birth cohort: Data from the Swedish twin registry

	Ages 50–60		Ages 60–70		Ages 70–80	
	Born 1922–1934	Born 1935–1947	Born 1901–1920	Born 1921–1938	Born 1890–1910	Born 1911–1929
Years of Schooling	0.933*	0.902**	0.972	0.957*	0.954	0.958**
	(0.872–0.999)	(0.841–0.968)	(0.923–1.023)	(0.920–0.995)	(0.891–1.022)	(0.928–0.989)
Number of Observations	9,505	14,432	10,683	12,925	5,365	12,017

Note: Confidence intervals are shown in parentheses.

*$p < .05$; **$p < .01$

Table 9 presents the mean expected remaining life expectancies at age 60, with the estimated scale and shape parameters (a,b) in parentheses. Low-educated men in the old cohort could expect to live an average of 19.1 more years at age 60, but highly educated men could expect to live an additional 3.5 years. The corresponding figures for females are 23.1 and 2.5, respectively. For the younger cohorts, the educational impact on expected survival is similar: at age 60, low-educated men could expect to live 20.7 more years, and highly educated men could expect to live 2.8 more years. Corresponding figures for females are 25.3 more years for low-educated females and 3 more years for highly educated females. Overall, it seems as if higher education is associated with 2.5–3.5 years additional life expectancy at age 60 compared with low education.

An alternative way to illustrate the impact of schooling on survival propensities is to study how the probability of reaching different ages is related to schooling. Contrary to the survival analysis based on the Gompertz distribution presented, this alternative allows for correction for unobserved within-twin-pair common characteristics. Table 10 presents twin fixed-effects estimates of linear probability models of survival up to ages 60, 70, and 80. The identification for these models emanates solely from twin pairs in which one twin has reached the considered age and the other one has not. Panel A of Table 10 presents results for the pooled sample of MZ and DZ twins.

The results shown in the first column of Table 10, panel A, imply that a highly educated person has a 4 percentage point greater chance of surviving to age 60 compared with a low-educated person. To put this in perspective, 94 % of low-educated persons survive until age 60. An increase in survival expectancy from 94 % to 98 % may not seem dramatic, but it is equivalent to a decrease in the mortality risk from 6 % to 2 % (or by almost 67 %). The numbers are slightly greater for males than for females.

The gender discrepancy is reversed for the chance of living until age 70: the advantage for the highly educated over the low educated is 8 percentage points among women and 6 percentage points among men. At age 80, the gender pattern is reversed again and the estimates again suggest a sizable reduction in the mortality risk for highly

Table 9 Remaining life expectancy at age 60 by birth cohort, educational group, and sex based on Gompertz estimations of the survival function: Data from the Swedish twin registry

	Born 1900–1919		Born 1920–1929	
Education	Males	Females	Males	Females
Low (<10)	19.112	23.125	20.740	25.317
	(0.0119, 10.354)	(0.0059, 9.164)	(0.0105, 10.855)	(0.0049, 9.617)
Middle (10–12)	20.269	24.920	21.763	25.798
	(0.0106, 10.529)	(0.0046, 9.086)	(0.0090, 10.641)	(0.0025, 7.389)
High (≥13)	22.621	25.613	23.575	28.302
	(0.0074, 11.044)	(0.0026, 7.347)	(0.0061, 9.586)	(0.0024, 8.247)

Notes: Estimated scale and shape parameters (a,b) are shown in parentheses. Estimates are based on Gompertz survival models for different subpopulations according to gender, educational group, and birth cohort, with birth year as an explanatory variable.

Table 10 Twin fixed-effects estimates of the probability of survival at ages 60, 70, and 80 by education and sex: Data from the Swedish twin registry

Education	Age 60 (birth year <1950)			Age 70 (birth year <1940)			Age 80 (birth year <1930)		
	Males and Females	Males	Females	Males and Females	Males	Females	Males and Females	Males	Females
A. MZ and DZ Twins									
Middle (10–12)	0.0148**	0.0140*	0.0154**	0.0215**	0.0232†	0.0197†	0.0239†	0.0324	0.0162
	(0.00418)	(0.00668)	(0.00523)	(0.00830)	(0.0133)	(0.0104)	(0.0138)	(0.0205)	(0.0186)
High (≥13)	0.0407**	0.0431**	0.0381**	0.0693**	0.0572*	0.0835**	0.108**	0.133**	0.0807†
	(0.00681)	(0.0105)	(0.00880)	(0.0154)	(0.0234)	(0.0205)	(0.0294)	(0.0422)	(0.0414)
Number of observations	40,170	18,410	21,760	28,462	12,848	15,614	20,802	9,292	11,510
B. MZ Twins									
Middle (10–12)	0.0230**	0.0370**	0.0123	0.0307*	0.0423†	0.0206	0.0356	0.0651†	0.0103
	(0.00692)	(0.0115)	(0.00839)	(0.0140)	(0.0229)	(0.0171)	(0.0233)	(0.0355)	(0.0309)
High (≥13)	0.0411**	0.0675**	0.0179	0.0607*	0.0444	0.0798*	0.116*	0.112	0.125†
	(0.0115)	(0.0184)	(0.0144)	(0.0266)	(0.0420)	(0.0340)	(0.0497)	(0.0744)	(0.0669)
Number of observations	14,920	6,824	8,096	10,246	4,634	5,612	7,390	3,336	4,054

Note: Standard errors are shown in parentheses.

†$p < .10$; *$p < .05$; **$p < .01$

educated persons. Turning to MZ twins, the precision of the estimates is lower because of smaller sample sizes. Nonetheless, for men and women combined as well as for women alone, the magnitude of the estimations closely mirrors those for the pooled MZ/DZ sample; for highly educated men, however, the estimates are about halved. Altogether, our estimates in this section suggest that the effect of schooling on mortality is large.

Comparisons With Previous Studies

Our fixed-effects results contrast with those of Behrman et al. (2011), who analyzed mortality, as measured by the propensities to survive until 2003 conditional on survival to 1980, among Danish twins born in 1921–1950. They found that the strong crude association between years of schooling and mortality (i.e., before 2003) vanished when the twins' common social and genetic background was accounted for. To assess whether this discrepancy between the findings of our study and those of Behrman et al. (2011) is due to differences in methodological approach, we replicate their study on MZ twins using the Swedish data. We include MZ twin pairs in which both were alive in 1980 and analyze the propensity to survive until 2003 in a linear twin fixed-effects probability model among the two cohorts born in 1921–1935 and 1936–1950 respectively. We include education as a continuous explanatory variable. Altogether, 10,290 Swedish twins meet these inclusion criteria.

For men in the Swedish data set, the resulting estimated effects of schooling on survival for both cohorts (0.0153 and 0.0080, respectively) are much greater than those Behrman et al. (2011) obtained (–0.0051 and 0.0012, respectively; both insignificant) and are statistically significant (at the 10 % and 5 % levels, respectively). If we instead use ages 67 and 52 as the endpoints for the respective cohorts, which is the maximum age attainable in 2003 for the youngest in the respective cohorts, the parameter estimates are similar in magnitude (0.0167 and 0.0064) and are statistically significant at the 5 % level. Thus, the differing results of the two studies are probably not a simple matter of varying endpoints of analysis. For women, the estimated effects, as in Behrman et al. (2011), become imprecise using their methodology.

What could explain the difference in results? Denmark and Sweden are neighboring countries that are both representatives of the Scandinavian welfare model. They also have similar economic and demographic characteristics. However, one noticeable difference between the two countries appears in life expectancy. Although life expectancy in both countries was among the highest in the world in 1950, life expectancy in Sweden has followed a rapidly rising pattern found in other Western countries during the last 50 years, whereas Denmark has fallen behind.[37] This divergence could be one cause of the difference between the two countries in the effects of schooling on longevity.[38]

[37] Around the millennium, the overall difference in life expectancy between the two countries was about three years, which has been attributed to a more unhealthy Danish lifestyle, especially in terms of alcohol and tobacco consumption (Juel 2008). Depending on its distribution within the population, this lifestyle may well interfere with the potential channels through which education may affect longevity.

[38] When we restrict our sample to the period 1921–1950 (the cohort studied by Behrman et al. (2011)) and run our SPL model on MZ and DZ twins, our estimate does not change much (0.950) and remains statistically significant.

Our results mirror those that Lager and Torssander (2012) obtained in their IV study using a schooling reform implemented in 1949–1962 in Sweden. They found a significant effect of the reform on mortality after the age of 40. Fischer et al. (2013) similarly found a small but positive effect on longevity of a 1936 Swedish schooling reform making the 7th year in school compulsory.

In our twin analysis, we exploit variation within twin pairs, which means that both twins faced the same rules regarding mandatory years of schooling, given that almost all twin pairs grew up in the same municipality. The variation we exploit therefore excludes exactly the variation exploited in the IV studies. Moreover, we have few twin pairs in which one twin has 7 (or 8) years of schooling and the other twin has 9 years of schooling. We find only 767 twin pairs in our data in which both twins have less than 10 years of schooling and differ in schooling. To obtain a somewhat comparable sample to the one used for identification in the IV studies, we can therefore restrict our sample to these 767 pairs and rerun our analysis. Doing so, we obtain an insignificant hazard ratio that is close to 1 (0.996). We thus obtain results quite similar to those of the IV studies when we restrict our sample to theirs. This finding suggests that our significant estimates in the main sample reflect variation at higher levels of schooling; it also makes the difference in results between our study and theirs less surprising.[39]

Summary and Conclusion

It is well known that well-educated people live longer. Less is known about why this is the case. In this article, we provide evidence on the topic by analyzing a large data set on Swedish twins born between 1886 and 1958. A unique feature of these data is that they enable us to observe complete life spans for a large number of individuals. By applying a twin fixed-effects design, we can account for many, often unmeasured, confounders, such as genes and family background. Our results show a strong and significant relationship between schooling and mortality, even when we account for such potential confounders. These results are robust to a number of controls for within-twin differences in early-life factors, such as birth weight and proxies for cognitive ability, and are relatively stable over the period considered.

Taken literally, our estimates suggest that the effect of schooling may be large; at age 60, the highly educated (with at least 13 years of schooling) could expect to outlive those with low education (less than 10 years) by an average of about 3 years. This is comparable to the current (2008–2012) gender discrepancy in life expectancy at age 60 in Sweden, where women expect to outlive men by about 3 years (Statistics Sweden 2014). It is also comparable to the gain in life expectancy (about 4 years) from kidney transplants among people aged 60 or older who are on dialysis (Wolfe et al. 1999), and the expected gains from quitting smoking in the United States at age 65 of about 2 years for men and 3 years for women (Taylor et al. 2002). However, the actual

[39] Another difference is that we are able to follow individuals until death in many cases, whereas they followed up until late adulthood. However, this difference seems to play a less important role because we find that the relationship between education and mortality becomes weaker with age rather than stronger.

causal effect of increasing schooling on mortality and life expectancy in contemporary Sweden is uncertain.

Our data go back to the end of the nineteenth century, when levels of schooling were considerably lower than today; the average years of schooling increased from about 9 years to 12 years during the period we study. Thus, we do not know if our results generalize to a situation with much higher levels of average schooling, as in contemporary Sweden. Clearly, however, the levels of schooling in our study period are similar to those found in many developing countries today. Our results therefore suggest that the longevity gains from extending levels of schooling in these countries may be substantial. Such gains and improved health may well constitute important mechanisms by which schooling also affects productivity and economic growth, which would further emphasize the crucial role of schooling in the development of human welfare.

From this perspective, the association between educational attainment (and income) and health/longevity has become stronger during the last decades in the United States, Europe, and Sweden, implying that the potential true societal value of schooling extends well beyond any economic growth effects (Hederos et al. 2014; Kunst et al. 2005; Masters et al. 2012; Miech et al. 2011). However, it is uncertain the extent to which this is a function of causal mechanisms (i.e., that education has become more important in promoting good health) or instead is a consequence of changes in the educational and income distributions (i.e., that the least-educated fraction of the population has become more negatively selected) (Montez and Friedman 2015).

Conceptualizing that mortality is an ultimate absorbing state of bad health, a closely related strand of the literature linking schooling to mortality addresses the association between schooling and health-related behavior and health. Interestingly, two recent twin-based studies from the United States and China find that the crude associations between schooling and health become much weaker when the twins' common backgrounds are taken into account (Amin et al. 2015a; Behrman et al. 2015). Mirroring these results, Gerdtham et al. (2016), analyzing a twin sample partly overlapping the one used in this study, found that the strong crude association between schooling and self-reported health is heavily reduced when common unobserved factors operating at a twin family level are taken into account. Hence, it seems as if there is a stronger connection between schooling and mortality than between schooling and (self-reported) health. The reasons for this difference are largely unknown, but a strong influence of schooling on mortality may not be associated with a similarly strong connection between schooling and survey-based (self-reported) health if the former effect is mainly due to causes of death from rather unexpected, quick-acting fatal illnesses.

Finally, some words of caution are warranted. Whereas the twin difference methodology we use neutralizes some of the influence from underlying, unobserved factors operating at a family level, it does not rule out the influence from environmental factors striking unevenly between the twins affecting both mortality and education (via cognitive ability, for example). We can take some of the influence from such unshared environmental factors into account, insofar as they are manifested by birth weight and adult height, but none of these entities are bound to capture the full lasting impact on schooling and health of, for example, illnesses and accidents affecting two twins unevenly during

childhood. Hence, our results clearly indicate that the association between schooling and mortality is not a consequence of environmental or genetic factors common to the twin pair or to divergent environmental influences as far as these are captured by birth weight and height. Although other underlying, unobserved, and unknown factors may affect both schooling and mortality, this implies that we are narrowing in on a causal interpretation with more transparent policy implications for the true value of investments in schooling.

We also caution against generalizing too much from country-specific studies. Comparing the results of this study with those of Behrman et al. (2011), a possible interpretation is that schooling is partly determined by the same endowments as mortality in Denmark (where no significant schooling effects were found when endowments were controlled for using twins fixed effects) but not in Sweden. Although the underlying reasons for it are uncertain, the discrepancy highlights the difficulties of generalizing from one context to another. If such generalizations are invalid even between relatively similar contexts, such as neighboring Denmark and Sweden, this has large implications for the risks of making global, or cross-continent generalizations from country-specific studies (e.g., between the United States and China).

In sum, the results of this study point to a strong connection between schooling and mortality. They also raise additional questions that need to be addressed in order to improve the understanding of how schooling may be a vehicle of improving longevity and health, as well as equality along these dimensions. First, the more precise nature of the underlying causal mechanisms behind the strong association between schooling and mortality found in this study remains largely unknown. Second, the puzzling weaker causal connection between schooling and (self-reported) health needs to be resolved. Previous population-wide research, partly based on schooling reforms, has demonstrated stronger correlations between education and mortality from preventable causes, such as alcohol abuse, lung cancer, and accidents (Lager and Torssander 2012; Mackenbach et al. 2015). Studying cardiovascular and ischemic heart disease, which are at least to some extent preventable, Madsen et al. (2014), found no significant association with education differences among Danish twins. Additional knowledge may be gained by analysis of specific causes of death, especially in a twin difference setting in which health, health-related behavior, and cause-specific mortality are jointly analyzed in order to establish the extent to which the association between schooling and mortality could be traced to preventable or unpreventable illnesses.

Acknowledgments We thank participants at the 22nd European Workshop on Econometrics and Health Economics and at the 23rd EALE conference for useful comments. We are also grateful for comments by seminar participants at Lund University and Linneaus University.

References

Adams, P., Hurd, M. D., McFadden, D., Merrill, A., & Ribeiro, T. (2003). Healthy, wealthy, and wise? Tests for direct causal paths between health and socioeconomic status. *Journal of Econometrics, 112,* 3–56.

Ahlstrom, K., Emanuelsson, I., & Wallin, E. (1986). *Skolans krav - elevernas behov* [School demands - student needs]. Lund, Sweden: Studentlitteratur.

Albouy, V., & Lequien, L. (2009). Does compulsory education lower mortality? *Journal of Health Economics, 28,* 155–168.

Almond, D., & Currie, J. (2011). Human capital development before age five. In D. Card & O. Ashenfelter (Eds.), *Handbook of labor economics* (Vol. 4B, pp. 1314–1486). Amsterdam, The Netherlands: North Holland.

Amin, V. (2011). Returns to education: Evidence from UK twins: Comment. *American Economic Review, 101,* 1629–1635.

Amin, V., Behrman, J. R., & Kohler, H.-P. (2015a). Schooling has smaller or insignificant effects on adult health in the US than suggested by cross-sectional associations: New estimates using relatively large samples of identical twins. *Social Science & Medicine, 127,* 181–189.

Amin, V., Behrman, J. R., Kohler, H.-P., Xiong, Y., & Zhang, J. (2015b). Causal inferences: Identical twins help and clarity about necessary assumptions is critical. *Social Science & Medicine, 127,* 201–202.

Amin, V., Behrman, J., & Spector, T. D. (2013). Does more schooling improve health behaviours and health outcomes? Evidence from U.K. twins. *Economics of Education Review, 35,* 134–148.

Ashenfelter, O., & Rouse, C. (1998). Income, schooling and ability: Evidence from a new sample of identical twins. *Quarterly Journal of Economics, 113,* 253–284.

Batty, G. D., Deary, I. J., & Gottfredson, L. S. (2007). Premorbid (early life) IQ and later mortality risk: Systematic review. *Annals of Epidemiology, 17,* 278–288.

Beauchamp, J. P., Cesarini, D., Johannesson, M., Lindqvist, E., & Apicella, C. (2011). On the sources of the height-intelligence correlation: New insights from a bivariate ACE model with assortative mating. *Behavior Genetics, 41,* 242–252.

Behrman, J. R., Kohler, H.-P., Jensen, V. M., Pedersen, D., Petersen, I., Bingley, P., & Christensen, K. (2011). Does more schooling reduce hospitalization and delay mortality? New evidence based on Danish twins. *Demography, 48,* 1347–1375.

Behrman, J. R., & Rosenzweig, M. R. (2004). Returns to birthweight. *Review of Economics and Statistics, 86,* 586–601.

Behrman, J. R., Rosenzweig, M. R., & Taubman, P. (1994). Endowments and the allocation of schooling in the family and in the marriage market: The twins experiment. *Journal of Political Economy, 102,* 1131–1174.

Behrman, J. R., Xiong, Y., & Zhang, J. (2015). Cross-sectional schooling-health associations misrepresented causal schooling effects on adult health and health-related behaviors: Evidence from the Chinese Adults Twins Survey. *Social Science & Medicine, 127,* 190–197.

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2007). From the cradle to the labor market? The effect of birth weight on adult outcomes. *Quarterly Journal of Economics, 122,* 409–439.

Boardman, J. D., & Fletcher, J. M. (2015). To cause or not to cause? That is the question, but identical twins might not have all of the answers. *Social Science & Medicine, 127,* 198–200.

Bonjour, D., Cherkas, L., Haskel, J., Hawkes, D., & Spector, T. (2003). Returns to education: Evidence from UK twins. *American Economic Review, 93,* 1799–1812.

Bound, J., & Solon, G. (1999). Double trouble: On the value of twin-based estimation of the return to schooling. *Economics of Education Review, 18,* 169–182.

Carlsson, M., Dahl, G. B., Ockert, B., & Rooth, D.-O. (2013). The effect of schooling on cognitive skills. *Review of Economics and Statistics, 97,* 533–547.

Case, A., & Paxson, C. (2008). Stature and status: Height, ability, and labor market outcomes. *Journal of Political Economy, 116,* 499–532.

Cesarini, D., Dawes, C. T., Johannesson, M., Lichtenstein, P., & Wallace, B. (2009). Genetic variation in preferences for giving and risk-taking. *Quarterly Journal of Economics, 124,* 809–842.

Clark, D., & Royer, H. (2012). The effect of education on adult mortality and health: Evidence from Britain. *American Economic Review, 103,* 2087–2120.

Cunha, F., & Heckman, J. (2007). The technology of skill formation. *American Economic Review, 97,* 31–47.

Currie, J. (2009). Healthy, wealthy, and wise? Socioeconomic status, poor health in childhood, and human capital development. *Journal of Economic Literature, 47,* 87–122.

Currie, J., Stabile, M. P., & Roos, L. (2010). Child health and young adult outcomes. *Journal of Human Resources, 45,* 517–548.

Cutler, D., Deaton, A., & Lleras-Muney, A. (2006). The determinants of mortality. *Journal of Economic Perspectives, 20,* 97–120.

Cutler, D. M., & Lleras-Muney, A. (2010). Understanding differences in health behaviors by education. *Journal of Health Economics, 29,* 1–28.

Cutler, D. M., & Lleras-Muney, A. (2012). *Education and health: Insights from international comparisons* (NBER Working Paper No. 17738). Cambridge, MA: National Bureau of Economic Research.

Fischer, M., Karlsson, M., & Nilsson, T. (2013). Effects of compulsory schooling on mortality: Evidence from Sweden. *International Journal of Environmental Research and Public Health, 10,* 3596–3618.

Fuchs, V. R. (1982). Time preference and health: An exploratory study. In V. R. Fuchs (Ed.), *Economic aspects of health* (pp. 93–120). Chicago, IL: University of Chicago Press.

Fujiwara, T., & Kawachi, I. (2009). Is education causally related to better health? A twin fixed-effect study in the USA. *International Journal of Epidemiology, 38,* 1310–1322.

Gathmann, C., Jtirges, H., & Reinhold, S. (2015). Compulsory schooling reforms, education and mortality in twentieth century Europe. *Social Science & Medicine, 127,* 74–82.

Gerdtham, U.-G., Lundborg, P., Lyttkens, C. H., & Nystedt, P. (2016). Do education and income really explain inequalities in health? Applying a twin design. *Scandinavian Journal of Economics, 118,* 25–48.

Glied, S., & Lleras-Muney, A. (2008). Technological innovation and inequality in health. *Demography, 45,* 741–761.

Grambsch, P. M., & Therneau, T. M. (1994). Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika, 81,* 515–526.

Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy, 80,* 223–255.

Hederos Ericksson, K., Jantti, M., Lindahl, L., & Torssander, J. (2014). *Trends in life expectancy by income and the role of specific causes of death* (SOFI Working Paper No. 8/2014). Stockholm: Swedish Institute for Social Research.

Holmlund, H., Lindahl, M., & Plug, E. (2011). The causal effect of parent's schooling on children's schooling: A comparison of estimation methods. *Journal of Economic Literature, 49,* 615–651.

Holsinger, T., Helms, M., & Plassman, B. (2007). Intelligence in early adulthood and life span up to 65 years later in male elderly twins. *Age and Ageing, 36,* 286–291.

Jang, K. L., Livesley, W. J., & Vemon, P. A. (1996). Heritability of the Big Five personality dimensions and their facets: A twin study. *Journal of Personality, 64,* 577–591.

Jayachandran, S., & Lleras-Muney, A. (2009). Life expectancy and human capital investments: Evidence from maternal mortality declines. *Quarterly Journal of Economics, 124,* 349–397.

Juel, K. (2008). Middellevetid og dpdelighed i Danmark sammenlignet med i Sverige: Hvad betyder rygning og alkohol? [Life expectancy and mortality in Denmark compared to Sweden. What is the effect of smoking and alcohol?]. *Ugeskrift for Laeger, 170,* 2423–2427.

Keller, M. C., Garver-Apgar, C. E., Wright, M. J., Martin, N. G., Corley, R. P., Stallings, M. C., . . . Zietsch, B. P. (2013). The genetic correlation between height and IQ: Shared genes or assortative mating? *PLoS Genetics, 9*(4). doi:10.1371/journal.pgen.1003451

Kenkel, D. (1991). Health behavior, health knowledge and schooling. *Journal of Political Economy, 99,* 287–305.

Kohler, H.-P., Behrman, J. R., & Schnittker, J. (2011). Social science methods for twins data: Integrating causality, endowments and heritability. *Journal of Biodemography and Social Biology, 57,* 88–141.

Kunst, A. E., Bos, V., Lahelma, E., Bartley, M., Lissau, I., Regidor, E., . . . Mackenbach, J. P. (2005). Trends in socioeconomic inequalities in self-assessed health in 10 European countries. *International Journal of Epidemiology, 34,* 295–305.

Lager, A., Bremberg, S., & Vagero, D. (2009). The association of early IQ and education with mortality: 65 year longitudinal study in Malmo, Sweden. *BMJ, 339,* 52–82.

Lager, A., & Torssander, J. (2012). Causal effect of education on mortality in a quasi-experiment on 1.2 million Swedes. *Proceedings of the National Academy of Sciences, 109,* 8461–8466.

Lichtenstein, P., De Faire, U., Floderus, B., Svartengren, M., Svedberg, P., & Pedersen, N. L. (2002). The Swedish 2007-04-27 21 Twin Registry: A unique resource for clinical, epidemiological and genetic studies. *Journal of Internal Medicine, 252,* 184–205.

Lleras-Muney, A. (2005). The relationship between education and adult mortality. *Review of Economic Studies, 72,* 189–221.

Lochner, L. (2011). Nonproduction benefits of education: Crime, health, and good citizenship. In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the economics of education* (Vol. 4, pp. 183–282). Amsterdam, The Netherlands: Elsevier Science.

Lundborg, P. (2013). The health returns to education—What can we learn from twins? *Journal of Population Economics, 26,* 673–701.

Lundborg, P., Nilsson, A., & Rooth, D. O. (2011). *Early life health and adult earnings: Evidence from a large sample of siblings and twins* (IZA Discussion Paper No. 5804). Bonn, Germany: Institute for the Study of Labor.

Lundborg, P., Nystedt, P., & Rooth, D. (2014). Height and earnings: The role of cognitive and non-cognitive skills. *Journal of Human Resources, 49,* 141–166.

Mackenbach, J. P., Kulhanova, I., Bopp, M., Deboosere, P., Eikemo, T. A., Hoffmann, R., . . . Lundberg, O. (2015). Variations in the relation between education and cause-specific mortality in 19 European populations: A test of the "fundamental causes" theory of social inequalities in health. *Social Science & Medicine, 127,* 51–62.

Madsen, M., Andersen, P. K., Gerster, M., Andersen, A. M. N., Christensen, K., & Osler, M. (2014). Are the educational differences in incidence of cardiovascular disease explained by underlying familial factors? A twin study. *Social Science & Medicine, 118,* 182–190.

Masters, R. K., Hummer, R. A., & Powers, D. A. (2012). Educational differences in U.S. adult mortality: A cohort perspective. *American Sociological Review, 77,* 548–572.

Mazumder, B. (2008). Does education improve health? A reexamination of the evidence from compulsory schooling laws. *Economic Perspectives, 32,* 2–16.

Meghir, C., Palme, M., & Simeonova, E. (2013). *Education, cognition and health: Evidence from a social experiment* (NBER Working Paper No. 19002). Cambridge, MA: National Bureau of Economic Research.

Miech, R., Pampel, F., Kim, J., & Rogers, R. G. (2011). The enduring association between education and mortality: The role of widening and narrowing disparities. *American Sociological Review, 76,* 913–934.

Miller, P., Mulvey, C., & Martin, N. (2005). Birth weight and schooling and earnings: Estimates from a sample of twins. *Economics Letters, 86,* 387–392.

Mirowsky, J., & Ross, C. E. (2003). *Education, social status, and health*. New York, NY: Aldine de Gruyter.

Montez, J. K., & Friedman, E. M. (2015). Educational attainment and adult health: Under what conditions is the association causal? *Social Science & Medicine, 127,* 1–7.

Nilsson, A. (2004). *Immigration and emigration in the postwar period* (Demographic Reports 2004:5). Stockholm: Statistics Sweden.

Oreopoulos, P., Stabile, M., Walld, R., & Roos, L. L. (2008). Short-, medium-, and long-term consequences of poor infant health: An analysis using siblings and twins. *Journal of Human Resources, 43,* 88–138.

Persico, N., Postlewaite, A., & Silverman, D. (2004). The effect of adolescent experience on labor market outcomes: The case of height. *Journal of Political Economy, 112,* 1019–1053.

Petersen, I., Jensen, V. M., McGue, M., Bingley, P., & Christensen, K. (2009). No evidence of genetic mediation in the association between birthweight and academic performance in 2,413 Danish adolescent twin pairs. *Twin Research and Human Genetics, 12,* 564–572.

Ridder, G., & Tunali, I. (1999). Stratified partial likelihood estimation. *Journal of Econometrics, 92,* 193–232.

Rosenzweig, M. R., & Schultz, T. P. (1982). The behavior of mothers as inputs to child health: The determinants of birth weight, gestation, and rate of fetal growth. In V. R. Fuchs (Ed.), *Economic aspects of health* (pp. 53–92). Chicago, IL: University of Chicago Press for the National Bureau of Economic Research.

Royer, H. (2009). Separated at girth: US twin estimates of the effects of birth weight. *American Economic Journal: Applied Economics, 1,* 49–85.

Sandewall, O., Cesarini, D., & Johannesson, M. (2014). Co-twin methodology and the return to schooling—Testing a critical assumption. *Labour Economics, 26,* 1–10.

Statistics Sweden. (2014). *Statistical yearbook of Sweden*. Orebro: Statistics Sweden, Communication Department.

Taylor, D. H., Jr., Hasselblad, V., Henley, S. J., Thun, M. J., & Sloan, F. A. (2002). Benefits of smoking cessation for longevity. *American Journal of Public Health, 92,* 990–996.

Torssander, J., & Erikson, R. (2010). Stratification and mortality—A comparison of education, class, status, and income. *European Sociological Review, 26,* 465–474.

van Kippersluis, H., O'Donnell, O., & van Doorslaer, E. (2011). Long run returns to education: Does schooling lead to an extended old age? *Journal of Human Resources, 46,* 695–721.

Webbink, D., Martin, N. G., & Visscher, P. M. (2010). Does education reduce the probability of being overweight? *Journal of Health Economics, 29,* 29–38.

Wolfe, R. A., Ashby, V. B., Milford, E. L., Ojo, A. O., Ettenger, R. E., Agodoa, L. Y., . . . Port, F. K. (1999). Comparison of mortality in all patients on dialysis, patients on dialysis awaiting transplantation, and recipients of a first cadaveric transplant. *New England Journal of Medicine, 341,* 1725–1730.