The relationship between health and schooling:  
What’s new?  

MICHAEL GROSSMAN 1,* 

1 City University of New York Graduate Center, National Bureau of Economic Research, and IZA

Abstract: Many studies suggest that years of formal schooling completed is the most important correlate of good health. There is much less consensus as to whether this correlation reflects causality from more schooling to better health. The relationship may be traced in part to reverse causality and may also reflect “omitted third variables” that cause health and schooling to vary in the same direction. The past five years (2010-2014) have witnessed the development of a large literature focusing on the issue just raised. I deal with that literature and what can be learned from it in this paper. I conclude that there is enough conflicting evidence in the studies that I have reviewed to warrant more research on the question of whether more schooling does in fact cause better health outcomes.

JEL classification: I10, I20

Key words: schooling, health, causality, efficiency, time preference

1 Introduction

For theoretical and empirical reasons, the positive relationship between more schooling and better health is one of the most fundamental ones in health economics. Clearly, the relationship is part of the massive literature in health economics on the determinants of the health of the population, a literature that originates from a demand for health model that I developed (for example Grossman, 1972a, 1972b, 2000). That model emphasizes that medical care is only one of many determinants of health, and it is natural to explore others. Moreover, my model views health as a form of human capital and therefore a determinant of earnings. Hence, it is natural to allow for and explore complementarities between health capital and other forms of human capital, the most important of which is knowledge capital, as proxied by the number of years of formal schooling completed.

Empirically, the importance of the relationship is highlighted by Kolata (2007, p. 1): “The one social factor that researchers agree is consistently linked to longer lives in every country where it has been studied is education. It is more important than race; it obliterates any effects of income.” It also is underscored by Meara et al. (2008, p. 350): “With the exception of black males, all recent gains in life expectancy at age twenty-five have occurred among better educated groups, raising education differentials in life expectancy by 30 percent.”

Table 1 contains additional empirical support of the importance of this relationship. The table shows trends in the health and educational attainment of the U.S. population between 1910 and 2000. These trends highlight the dramatic improvements in these outcomes in the past century. To be specific, the infant mortality rate fell by a factor of almost twenty, and the age-adjusted mortality rate declined by a factor of almost three. At
the same time, there was almost a ten-fold increase in the percentage of the population who completed four years of college or more.

Table 1: Infant mortality rate, age-adjusted mortality rate, and educational attainment, United States, selected years, 1910-2000

<table>
<thead>
<tr>
<th>Year</th>
<th>Infant mortality rate (deaths per 1,000 live births)</th>
<th>Age-adjusted mortality rate (deaths per 100,000 population based on year 2000 standard population)</th>
<th>College graduates (percentage of persons aged 25 and older who completed four years of college or more)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1910</td>
<td>131.8</td>
<td>2,317.2</td>
<td>2.7</td>
</tr>
<tr>
<td>1920</td>
<td>92.2</td>
<td>2,147.1</td>
<td>3.3</td>
</tr>
<tr>
<td>1930</td>
<td>69.0</td>
<td>1,943.8</td>
<td>3.9</td>
</tr>
<tr>
<td>1940</td>
<td>54.9</td>
<td>1,785.0</td>
<td>4.6</td>
</tr>
<tr>
<td>1950</td>
<td>33.0</td>
<td>1,446.0</td>
<td>6.2</td>
</tr>
<tr>
<td>1960</td>
<td>27.0</td>
<td>1,339.2</td>
<td>7.7</td>
</tr>
<tr>
<td>1970</td>
<td>21.4</td>
<td>1,222.6</td>
<td>11.0</td>
</tr>
<tr>
<td>1980</td>
<td>12.9</td>
<td>1,039.1</td>
<td>17.0</td>
</tr>
<tr>
<td>1990</td>
<td>9.7</td>
<td>938.7</td>
<td>20.3</td>
</tr>
<tr>
<td>2000</td>
<td>7.4</td>
<td>869.0</td>
<td>25.6</td>
</tr>
</tbody>
</table>

Total change, 1910-2000a 124.4 1,448.2 22.9


aAbsolute value.

To summarize the trends in Table 1, I have run regressions of each of the two mortality rates on the schooling variable and trend terms. Given the high correlation between schooling and time, the estimation of these regressions is perhaps more of an art than a science. In general, my procedure is to experiment with linear, quadratic, and cubic trend specification and to select the one with the lowest residual variance. For infant mortality, the cubic model outperforms the other two. For age-adjusted mortality, the quadratic and cubic models do about the same. To be consistent, I select the cubic model for both outcomes. The schooling coefficient in the age-adjusted mortality regression is not, however, sensitive to this selection.

The regression results are presented in Table 2. There are 50 observations in the infant mortality equation because the schooling variable was available in 1910, 1920, 1930, 1940, 1947, 1950, 1952, 1957, 1959, 1960, 1962, and 1964-2002 (the last year in which infant mortality was available when I compiled the series). There are 49 observations in the age-adjusted mortality equation because that series ended in 2001.
### Table 2: Infant and age-adjusted mortality regressions

<table>
<thead>
<tr>
<th>Percentage with four years of college or more</th>
<th>Infant mortality rate</th>
<th>Age-adjusted mortality rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.617</td>
<td>-1.617</td>
<td>-28.950</td>
</tr>
<tr>
<td>(-5.06)</td>
<td>(-3.96)</td>
<td></td>
</tr>
<tr>
<td>R²</td>
<td>0.996</td>
<td>0.990</td>
</tr>
<tr>
<td>F-statistic</td>
<td>2,814.71</td>
<td>1,078.82</td>
</tr>
</tbody>
</table>

Each regression contains an intercept and a cubic time trend. t-statistics are given in parentheses. The sample size is 50 in the first regression and 49 in the second regression. See text for more details.

The schooling coefficient is negative and statistically significant in each of the two regressions. The growth in schooling “explains” approximately 30 percent of the reduction in infant mortality between 1910 and 2000 and approximately 48 percent of the reduction in age-adjusted mortality.\(^1\) The regression results in Table 2, however, and the evidence provided by Kolata (2007) and by Meara et al. (2008) do not necessarily imply causality from more schooling to better health. Health may cause schooling or omitted “third variables” may cause health and schooling to vary in the same direction.

Students in poor health are almost certain to miss more days of school due to illness than their healthy peers and may also learn less while they are in school. Both factors suggest negative effects of poor health in childhood on school achievement and ultimately on years of formal schooling completed. Furthermore, this causal path may have long-lasting effects if past health is an input into current health status. Thus, even for non-students, a positive relationship between health and schooling may reflect causality from health to schooling in the absence of controls for past health. Health also may cause schooling because a reduction in mortality increases the number of periods over which the returns from investments in knowledge can be collected.

Productive and allocative efficiency models generate causality from education to health. In the former model, the more educated are assumed to obtain more health output from given amounts of medical care and other inputs. In the latter model, the more educated are assumed to pick a different input mix to produce health than the less educated. That mix gives them more output than the mix selected by the less educated.

Since health and schooling are both endogenous, unobserved “third variables” may cause both of these outcomes to vary in the same direction. Fuchs (1982) identifies time preference as perhaps the key third variable. He argues that persons who are more future oriented (who have a high degree of time preference for the future or discount it at a modest rate) attend school for longer periods of time and make larger investments in their own health and in the health of their children. Thus, the effects of schooling on these outcomes are biased if one fails to control for time preference.\(^2\)

During the past five years (2010-2014), many studies have appeared that attempt to assess whether more schooling does in fact cause better health or health behaviors that contribute to this outcome. In the remainder of this paper, I focus on these studies and what can be learned from them.

---

1. For infant mortality, this was calculated as 100* \([22.9*1.617]/124.4\) = 30 percent. For age-adjusted mortality, the computation is 100*\([22.9*28.950]/1,448.2\) = 48 percent.
2. For a detailed discussion of the framework employed to study the health-schooling relationship, see Grossman (2006) and the references that I cite in that paper. My discussion includes a model in which schooling causes health because it makes people more future oriented.
2 Classification and inventory of studies

Table 3 contains a classification and inventory of the kinds of studies in the period from 2010 through 2014 that I will discuss. Thirty-eight studies are listed in the table, although there may be others with which I am not aware. These studies are of three types. The first type directly includes such hard-to-measure third variables as time preference, cognitive development, noncognitive development, and past health. The second type controls for unobserved genetic and environmental factors by examining the effects of differences in schooling obtained by identical twins on differences in their health outcomes. The third type employs the technique of instrumental variables. The idea here is to find exogenous variables that are correlated with schooling but not correlated with unmeasured variables that affect health. In addition, they have no effect on health, with schooling held constant. These variables serve as instruments for schooling in the estimation of health equations by two-stage least squares. One such instrument is the enactment of a law that increases the required amount of formal schooling.

Table 3: Classification and inventory of studies, 2010-2014

<table>
<thead>
<tr>
<th>Direct inclusion of third variables (8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conti and Hansman (2013); Conti and Heckman (2010); Conti et al. (2010); De Walque (2010); Kaestner and Callison (2011); Savelyev (2014); Savelyev and Tan (2014); Van Der Pol (2011)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Twin studies (9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amin and Behrman (2014); Amin et al. (2013); Amin et al. (2011); Behrman et al. (2011); Lundborg et al. (2012a); Lundborg (2013); Lundborg et al. (2012b); Madsen et al. (2010); Webbink et al. (2010)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Instrumental variables studies (21)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agüero and Bharadwaj (2013); Andalón et al. (2014); Atella and Kopinska (2014); Braakmann (2011); Buckles et al. (2013); Carneiro et al. (2013); Chou et al. (2010); Clark and Royer (2013); Cowan (2011); Dinçer et al. (2014); Fabrice and Jones (2011); Jensen and Lleras-Muney (2012); Jürges et al. (2011); Kemptner et al. (2011); Lundborg et al. (2014); McCrary and Royer (2011); Meghir et al. (2012); Mocan and Cannonier (2012); Powdthavee (2010); Van den Berg et al. (2012); Van Kippersluis et al. (2011)</td>
</tr>
</tbody>
</table>

Many of the studies in Table 3 employ a simple conceptual and econometric framework. Contributions by Nobel Laureate James J. Heckman and his colleagues (Conti and Heckman, 2010; Conti et al., 2010; Savelyev, 2014; Savelyev and Tan, 2014) are of a very different nature. They employ the framework developed by Heckman and colleagues (for example, Cuhna et al., 2010), one that conceptualizes the technology of cognitive and noncognitive skill formation in childhood and adolescence as key determinants of completed schooling and health as an adult. In their framework, self-productivity (skills produced early in life increase skills at later stages) and dynamic complementarity (early investments raise the marginal product of later investments) interact to generate multiplier effects. Hence, investments in adolescence have much larger payoffs when earlier investments are made.

Heckman and colleagues stress the importance of noncognitive skills as key third variables. They include in these skills the “big five personality traits”: conscientiousness, openness, extraversion, agreeableness, and neuroticism. Their econometric framework is much more complicated than that in the other studies that I will discuss. It incorporates
latent cognitive and personality skills, measured and unmeasured components of those skills, factor analysis to control for measured and unmeasured cognitive development, and the simultaneous estimation of outcome and measurement equations.

The studies in Table 3 include research that addresses (1) the relationship between an individual’s own schooling and his or her own health; (2) the relationship between parents’ schooling and their children’s health; and (3) the relationship between schooling and mechanisms that may lead to worse or better health outcomes. Examples are health knowledge, fertility choices, and such unhealthy behaviors as cigarette smoking, excessive alcohol consumption, and overeating and lack of exercise – sometimes reflected by a large body mass index (BMI) and obesity. In the remainder of this paper, I will discuss empirical results in the studies in Table 3. Given the large number of studies in that table, I cannot deal with all of them. I do try, however, to give a balanced picture of the nature of the findings.

3 Inclusion of third variables

The eight studies in Table 3 that include hard-to-measure third variables all find positive and significant effects of completed schooling on at least some key measures of adult health and beneficial health behaviors. For example, Van Der Pol (2011) controls for time preference in the Dutch DNB Household Survey. In that survey, respondents were asked how much they would willing to give up today in order to get a certain amount of money next year. Outcomes include self-rated health, cigarette smoking, body mass index (BMI), and obesity.

To cite another example, Conti and Heckman (2010) control for cognitive and non-cognitive ability at age 10 and health at that age in an examination of outcomes at age 30 in the 1970 British Cohort Study. These outcomes include self-rated health, daily smoking, and obesity. Conti and Heckman report positive effects of education on self-rated health and the negative effects of this variable on smoking and obesity. The education effects are more important for those with higher levels of cognitive ability, but less important for those with higher levels of noncognitive ability.

To cite a final example, Savelyev (2014) focuses on life expectancy in the Terman Life Cycle Study of Children with High Ability. Members of this sample have IQs greater than 140, which corresponds to the 99.6th percentile of the IQ distribution. They were 11 years old in 1921, were followed through 1991, and all were high school graduates. For men, graduation from college increases life expectancy at age 30 by approximately 9 years compared to non-graduates. This result holds the big five personality traits and health status (all measured at age 12) constant. No effects are observed for women. The findings for men indicate that the favorable effects of education on health persist among those with extremely high levels of cognitive ability and are not limited to completion of primary or secondary school.

How much do the significant schooling effects fall once third variables are held constant? Given differences in outcomes, third variables, and samples in the three studies just reviewed and in others, there is no straightforward answer to this question. Van der Pol (2011) reports that the effects of schooling on self-rated health falls between 7 and 14 percent in alternative specifications once time preference is held constant. The inclusion of the latter variable has, however, no impact on the effects of schooling on smoking, BMI, and obesity. Conti and Heckman, who find that the education effect is largest in the case of daily smoking, indicate a reduction of approximately 25 percent in this effect when

I consider fertility as an outcome because children in large families tend to have worse health outcomes than those in smaller families.
cognitive and noncognitive skills and health at age 10 are employed as regressors. Savelyev (2014) does not consider this issue. He does indicate, however, that the 9-year increase in life expectancy for college graduates compared to high school graduates mentioned above exceeds the 6-year differential in that outcome between the highest and the lowest deciles of his key third variable: conscientiousness. The important message in the studies that employ direct measures of third variables is that there are few instances in which large schooling effects are reduced by a significant percentage by these variables or in which the net effects of these variables exceed the net effects of schooling.

4 Twin differences

Results from studies that focus on twin differences have the flavor of a “point-counterpoint.” Behrman et al. (2011) find no effects of differences in education on differences in adult mortality or hospitalizations in a Danish twin registry that consists of 2,500 identical (monozygotic, MZ) twin pairs. Amin et al. (2013) report no effects of education on obesity, smoking, and physical health in a sample of 741 female MZ twin pairs in the United Kingdom. Madsen et al. (2010) indicate no overall effects in the same sample employed by Behrman et al. (2011) but negative effects in the outcomes considered by Behrman and colleagues for males born before 1935 and negative effects for cases in which there are large schooling differences within twin pairs.

The last two findings are counterpoints to the absence of schooling effects in the first two studies. Webbink et al. (2010) provide another counterpoint. They report a negative impact of schooling on male obesity in 350 Australian MZ twin pairs. They do not, however, find an effect in the 350 MZ female pairs.

In three studies, Lundborg and colleagues provide somewhat stronger counterpoints to the conclusion that more schooling does not cause better health in twin data (Lundborg, 2013; Lundborg et al., 2012a; Lundborg et al., 2012b). Lundborg (2013) reports effects in the expected direction of high school completion on self-rated health, chronic conditions, and exercise behavior among 347 MZ pairs in the Midlife in the United States survey. There are, however, no effects on smoking and BMI.

Lundborg et al. (2012a) find negative effects of schooling on mortality in a Swedish twin registry consisting of 9,000 MZ pairs. Individuals with at least 13 years of schooling can expect to live an additional 24 years at age 60 compared to 21 years for those with less than 10 years of schooling. Moreover, 84 percent of low-educated individuals lived to age 70, compared to 90 percent of high-educated individuals. These results control for twin differences in birthweight and height.

Lundborg et al. (2012b) report a negative effect of MZ twin mother’s schooling on her son’s health in the Swedish twin registry employed in the previous study. The finding pertains to sons who have enlisted in the military, with health measured by physical examination administered to 300 twin pairs on enlistment. There is no effect of MZ twin father’s schooling on the health outcome.

5 Instrumental variables

All but one of the 22 instrumental variables studies in Table 3 use compulsory school reform or school entry cutoff dates as the instruments for schooling, sometimes combined with new school openings at a differential rate among areas. These studies account for more than half of those in the table, and I will group my discussion of them by the outcomes they consider. First, I will consider those in which adult mortality is the outcome. Second, I will review those in which the outcome is adult health, health behaviors, or health knowledge. Finally, I will comment on those that focus on the impacts of parents’ schooling on infant and
adolescent health. Like the twin differences studies, the instrumental variables studies have the flavor of a point-counterpoint.

### 5.1 Adult mortality

Clark and Royer (2013) report no effect in Britain. They conclude (p. 2089): “Our results...suggest that economic models that assume a strong causal effect of education on health ought to be carefully reconsidered.” Meghir et al. (2012) present a similar lack of evidence of a schooling effect in Sweden. Van Kippersluis et al. (2011) indicate a very different finding for men but not for women in the Netherlands. For men surviving to age 81, a one-year increase in the number of years of formal schooling obtained due to compulsory school reform lowers the probability of dying before age 89 by 3 percentage points relative to a baseline percentage of those who will die in that period of 50 percent. Buckles et al. (2013) report a complementary finding for men in the United States. In their study, college completion reduces cumulative mortality by almost 30 percent relative to the mean for men ages 38-49 in 1980. Given the evidence in the latter two studies, the conclusion by Clark and Royer (2013) seems somewhat premature.

### 5.2 Adult health and health behaviors

When adult health and health behaviors are outcomes, the evidence also is mixed. Braakmann (2011) finds no effects on variety of self-rated health measures, smoking, heavy drinking, and diet in Britain. Similar findings are reported by Clark and Royer (2013). Powdthavee (2011) indicates a negative effect on hypertension based on physical exam in Britain. Fabrice and Jones (2011) indicate negative effects on smoking and quitting smoking in France. Buckles et al. (2013) find negative effects on smoking, heavy drinking, and obesity; and a positive effect on exercise in the U.S. Atella and Kopinska (2014) report negative effects on BMI and caloric intake and a positive effect on calorie expenditure in Italy.

### 5.3 Health knowledge and related behaviors

Agüero and Bharadwaj (2013) show that more educated women in Zimbabwe have more knowledge about HIV and fewer sexual partners than their less educated peers. Andalón et al. (2014) show that more educated Mexican women have more contraceptive knowledge and are more likely to use contraception at sexual debut than less educated women. Dinçer et al. (2014) report similar results for the effects of Turkish women’s education on their knowledge of the ovulation cycle and their use of modern family planning methods. Finally, Mocan and Cannonier (2012) find positive effects of women’s schooling on their use of modern contraception and on the likelihood that they are tested for AIDS in Sierra Leone.

### 5.4 Infant health, adolescent health, and mechanisms

McCrary and Royer (2011) fail to uncover effects of mother’s schooling on low birthweight, infant mortality, maternal smoking and alcohol use during pregnancy, and prenatal care use for pregnant women in California and Texas. Chou et al. (2010) uncover negative effects of mother’s schooling on low birthweight, neonatal mortality, postneonatal mortality, and infant mortality in Taiwan. An increase in schooling associated with compulsory school reform saved almost 1 infant life in 1,000 live births in that country. Lundborg et al. (2014) show that an increase mother’s schooling has positive effects on their sons’ physical health

---

4 Buckles et al. (2013) do not use compulsory school reform or school entry age laws as their instrument for schooling. Instead, they employ the risk of induction into the armed services during the Vietnam War era in that capacity.
and height in Swedish military enlistment register data. Dinçer et al. (2014) find positive effects of mother’s schooling on age at first marriage and at first birth, a negative effect on number of pregnancies, and weak evidence of negative effect on infant mortality in Turkey.

6 Discussion

There is enough conflicting evidence in the studies that I have reviewed to warrant more research on the question of whether more schooling does in fact cause better health outcomes. Future research in this area will be especially helpful if it confronts and tries to address the following issues:

- School entry ages used by McCrary and Royer (2011) and others as instrument may not be exogenous because parents can hold their children back or petition to have them start early. Shigeoka (2014) finds almost 2,000 births per year shifted from a week before to a week after the school entry cutoff date in Japan.

- Why do the results of instrumental variables studies vary so much? Is it because the instruments used in different studies may affect different groups, so that local average treatment effects, rather than average treatment effects, are being obtained?

- Twin studies typically have small sample sizes. Differences in schooling between identical twins may be small. Why do identical twins obtain different amounts of schooling?

- Some instrumental variables studies find that ordinary least squares estimates are consistent. Some do not test this. Schooling is endogenous, but is it possible that recursive model with uncorrelated errors is correct? That is, to what extent do third variables influence schooling but have no impact on health with schooling held constant?

- Studies I discussed focus on years of schooling completed; what about the quality of schooling? There is some work in this area (for example, Frisvold and Golberstein, 2011), but much more could be done.

- What are the mechanisms via which schooling influences health and health behaviors?

- Does it provide specific health information or general knowledge that leads to better health behaviors and improved health outcomes? Alternatively, is more schooling a vehicle in the construction of a future orientation, one that is required to make decisions today that will have favorable consequences for many years to come?

Acknowledgement

This paper is based on my keynote address at the Thirty-Fifth Annual Conference of the Nordic Health Economists’ Study Group at the University of Iceland in Reykjavik, Iceland, August 20-22, 2014. I am indebted to the organizers of the conference for inviting me to make the presentation. There are no conflicts of interest.
References


© 2015 the author(s). This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).