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

Vikesh Amin[a]
Jere R. Behrman[b],
Hans-Peter Kohler[c]

9 December 2013

Abstract: Adult health outcomes and health behaviors generally are strongly associated with schooling attainment. But such associations do not necessarily imply that schooling has causal effects on health outcomes and behaviors of the magnitudes of the associations. Schooling may be proxying for unobserved factors that are related to genetics and family background. Recently several studies have used within-identical (monozygotic, MZ) twins methods to control for those unobserved factors that are shared completely by identical twins. Estimates based on relatively small samples for the US, as well as some larger samples for other countries, suggest that causal impacts of schooling on health outcomes and behaviors are insignificant or much smaller than suggested by cross-sectional associations. This study contributes new estimates of cross-sectional associations and within-MZ causal effects of twins using three relatively large US samples: Mid-Atlantic Twin Registry, Minnesota Twin Registry and NAS-NRC Twin Registry of WWII Military Veterans. The estimates suggest that schooling is significantly associated with numerous health outcomes and behaviors in the US. However if within-MZ twins estimators are used to control for unobserved factors, there is no causal relationship between schooling and better health behaviors. There is some evidence that more schooling causally affects self-reported health and overweight status, net of unobservable cofounders, though not to the extent that cross-sectional associations suggest. Finally, spousal schooling is associated with better health outcomes and behaviors, but there is no evidence of any causal effect.

[a] Corresponding Author. Assistant Professor, Department of Economics, Central Michigan University, Mount Pleasant, Michigan, MI 48859, USA, amin1v@cmich.edu. Tel:1-989-774-3372. Fax: 1-989-774-2040
[b] William R. Kenan, Jr. Professor of Economics and Sociology, McNeil 160, 3718 Locust Walk, University of Pennsylvania, Philadelphia, PA 19104-6297, USA, jbehrman@econ.upenn.edu.
[c] Fredrick J. Warren Professor of Demography, McNeil 272, 3718 Locust Walk, University of Pennsylvania, Philadelphia, PA 19104-6297, USA, hpkohler@pop.upenn.edu
Acknowledgements: The National Institute of Child Health and Development grant RO1 HD046144-01 on "Causal Effects of Schooling on Adult and Child Health" provided background support for this study.

Introduction

It is well-documented that more-schooled individuals tend to have better health outcomes and health behaviors. Economic theory posits that a causal relationship may exist, because more-schooled individuals are more efficient in *health production* (Grossman 1972), that is, the process by which individuals use health inputs like time, money and behaviors to "produce" health over the life-course. The *productive efficiency theory* hypothesizes that more-schooled individuals produce more health from a given set of inputs. The *allocative efficiency hypothesis* states that schooling makes individuals aware of the negative effects of certain health behaviors and more efficient at using health information. More-schooled individuals also have higher incomes, and are able to afford better health care.

However, associations between schooling and health do not necessarily reflect causal relationships that imply that increased schooling *improve* individuals' health. Instead the associations could be cofounded by unobserved factors such as ability that jointly affect schooling and health. For example, Cutler and Lleras-Muney (2010) found that 20 percent of the schooling health gradient is driven by cognitive ability. Moreover, there is also possibly reverse causation, insofar as health (particularly early life health) affects schooling.

Recent studies have attempted to estimate causal effects of schooling in the US using instrumental variable (IV) estimates based on natural policy experiments regarding schooling. Some studies have instrumented college completion of men using variation in college attainment induced by draft avoidance behavior during the Vietnam War. These studies have found college completion reduced mortality, smoking, body mass index (BMI) and the probability of being overweight (De Walque 2007, Grimard and Parent 2007, MacInnis 2006, Buckles et al. 2012). Other studies have used changes in compulsory schooling laws as an instrument and found that more schooling improved self-reported health (Mazumder 2008), old age cognitive functioning (Glymour et al. 2008), reduced BMI and probabilities of being overweight (Grabner 2008), but had no effects on hospitalizations, functional limitations and specific health conditions (Mazumder 2008). Using changes in compulsory schooling laws Lleras-Muney (2005) found that more schooling reduced mortality, but Mazumder (2008) showed these results were not robust to inclusion of state-specific trends.

An alternative to the IV strategy using policy changes is to use data on MZ (monozygotic; identical) twins and within-MZ twin estimators. MZ twins share the same genetic makeup and family-rearing environment. By relating twin-pair differences in health measures to twin-pair differences in schooling, schooling effects on health can be estimated controlling for influences of unobserved factors that affect schooling and health of both twins. Fujiwara and Kawachi (2009) and Lundborg (2013) applied this strategy using data for 694-701 MZ twins from the 1995 Midlife in the United States Survey (MIDUS). Both studies found no significant effects of schooling on health outcomes (self-reported health, perceived mental health, number of chronic conditions, BMI) and health behaviors (smoking, exercise). Lundborg (2013) did find significant postive causal impacts of schooling on self-reported health and negative causal impacts on number of chronic conditions, when using dichotomous schooling indicators for high school, some college and college degree instead of a continous measure of grades of schooling. Other studies have been published on other countries. For example, Behrman et al. (2011) used this approach to investigate relations between schooling and hospitalization and mortality using data on 5,294 MZ twins in Denmark. They found significant cross-sectional associations, but no significant effects of schooling with within-MZ twin estimators.

This study provides more evidence of the schooling health-gradient using within-MZ twins methodology and much larger US twin datasets than in MIDUS from the Mid-Atlantic Twin Registry, Minnesota Twin Registry and the NAS-NRC Twin Registry of WWII Military Veterans. These new explorations with larger sample sizes permitted assessment of whether the mostly insignificant results reported in the above studies for the US were in part the artifact of relatively small sample sizes. We found that more schooling was associated with better health outcomes and behaviors. With control for unobserved factors, through within-MZ twins estimators, there was no causal relationship between schooling and better health behaviors. We found, however, some evidence that more schooling was associated with better health (self-reported health and overweight status), net of unobserved cofounders, though not to the extent that the cross-sectional associations suggested. Finally, we investigated whether there were positive spillover effects of spousal schooling on health. We found that spousal schooling was associated with better health outcomes and behaviors, but no evidence of any causal effects.

Datasets

The Mid-Atlantic Twin Registry (MATR)

The MATR is a population-based registry of twin pairs ascertained from birth records and school system records of Virginia, North Carolina, and South Carolina (see Anderson et al. 2002 for details). We used data from the Virginia 30K MATR sample. The sample was ascertained from public birth records in Virginia for twins born in 1915-1972 and responses to a letter published in the newsletter of the American Association of Retired Persons. Twins were mailed health and lifestyle questionnaires in 1987. Completed questionnaires were obtained from 14,763 twins, approximately 70 percent. Our analysis was based on samples of white MZ twin pairs aged 25-75, with non-missing information on self-reported schooling, reports on the co-twin's schooling, and health measures. The sample sizes ranged from 3,392 to 4,328.

Twins were asked to report their own schooling, their co-twin's schooling, and their spouse's schooling. Schooling attainment was measured using a 6 point scale: 0-7 grades of elementary school; 8 grades of elementary school; 1-3 grades of high school; 4 grades of high school; 1-3 years of college; 4 or more years of college. These categories were recoded as 5, 8, 10, 12, 14 and 16 schooling grades, respectively.

Self-reported health was measured by "how would you rate your health in the past 12 months" with responses ranging from 1=very poor, 2=poor, 3=average, 4=good and 5=very good. BMI was calculated as (weight in kg)/(height in meters)^2, where both height and weight were self-reported. A twin was defined as being overweight (obese) if his/her BMI was greater or equal to 25 (30). Never smoked was measured through an indicator equal to 1 if the twin had never smoked and 0 otherwise. Twins were asked how many alcoholic drinks they consumed in a typical week with options for 0, 1-3, 4-6, 7-12, 13-18, 19-24, 25-42 and more than 42 drinks per week. We recoded these categories as 0, 2, 5, 8.5, 15.5, 21.5, 33.5 and 53 respectively. Twins were asked to describe their leisure exercise with options for no exercise; occasional exercise; regular exercise about once a week; exercise a couple of time a week or jogging, cycling to work or vigorous sport activity at least 3-4 times a week. Exercise was measured through a dichotomous variable equal to 1 if the twin reported regular exercise once a week, exercise a couple of times a week or jogging, cycling to work or vigorous sport activity at least 3-4 times a week and 0 otherwise.

The Minnesota Twin Registry (MTR)

The MTR contains birth records on all twins born in Minnesota in 1936-1955 (details are in Lykken et al. 2002). We used data from the Socioeconomic Survey of Twins, which was a questionnaire sent in 1994 to 6638 members of same sex twin pairs. There was a 55 percent response rate, with 3680 valid questionnaires returned. Our analysis was based on samples of white MZ twin pairs, with non-missing information on self-reported schooling, co-twin's schooling, and health measures. The sample sizes ranged from 1,290 to 1,310.

Twins were asked to report their own schooling, their co-twin's schooling, and their spouse's schooling. The questionnaire contained several questions pertaining to schooling attainment: (1) highest grade 1 through 12 of schooling completed; (2) vocational schooling (3) college schooling (4) graduate/professional schooling and (5) schooling at time of first marriage and current marriage. We first determined the highest qualification obtained and assigned actual grades of schooling if no high school diploma, 11 if GED, 12 if high school diploma, 13 if vocational diploma, 14 if associate degree, 16 if college degree, 18 if masters degree, 19 if JD or MBA and 20 if doctoral degree.

The questionnaire contained a limited number of health questions. Twins were asked "how would you rate your health at the present time?" with responses for 1=bad, 2=poor, 3=fair, 4=good, 5=excellent. BMI was calculated as weight in kg/height in meters^2, where both height and weight were self-reported. Twins were defined as being overweight (obese) if their BMI was greater or equal to 25 (30). No information on health behaviors was available.

NAS-NRC Twin Registry of WWII Military Veterans

The NAS-NRC twin registry consists of 15,924 white male twins born in 1917-1927, both of whom served in the armed forces. Twins were mailed questionnaires in 1967, 1983 and 2000. We used data from the 2000 survey, which contained information on both health and schooling. 2060 twin pairs responded to the questionnaire (Page 2002). Our analysis was based on samples of white MZ twin pairs, with non-missing information on self-reported schooling, and health measures. The sample sizes ranged from 1726 to 1902.

The questionnaire asked twins their own schooling, but not their co-twins' or spouses' schooling. Schooling attainment was measured through a single question asking "what is your highest grade level, diploma, degree completed?", with options for none, 1 grade, 2 grades, 3 grades, 4 grades, 5 grades, 6 grades, 7 grades, 8 grades, 9 grades, 10 grades, 11 grades, high school or GED, 1 year of trade/vocational school after high school, 1 year of college, 2 years of college or associates degree, 3 years of college, bachelors degree, 1 year of graduate work, masters degree, some doctoral work, doctoral degree. We assigned actual grades if high school was not completed, 12 grades for graduating from high school, 13 grades for 1 year of trade school, 13 grades for 1 year of college, 14 grades for 2 years of college/associate degree, 15 grades for 3 years of college, 16 grades for a bachelors degree, 17 grades for 1 year of graduate work, 18 grades for a masters degree, 19 grades for some doctoral work and 20 grades for a doctoral degree.

Self-reported health was measured by "in general would you say your health is" 1=poor, 2=fair, 3=good, 4=very good, 5=excellent. BMI was calculated as weight in kg/height in meters^2, where both height and weight were self-reported. A twin was defined as being overweight (obese) if his/her BMI was greater or equal to 25 (30). Twins were asked whether their health limited (1) vigorous activities, (2) moderate activities, (3) lifting or carrying groceries, (4) climbing several flights of stairs, (5) climbing one flight of stairs, (6) bending, kneeling or stooping, (7) walking more than one mile, (8) walking several blocks, (9) walking one block and (10) bathing or dressing yourself. The questions had 3 possible responses- 1=no, not limited at all, 2=yes limited a little, 3=yes, limited a lot. We created an overall measure encapsulating how much health limited activities by summing responses to the 10 questions. For specific health conditions, we examined dichotomous variables indicating whether twins have ever had a heart attack, a stroke, diabetes, high blood pressure, prostate cancer. As a final health outcome, we examined the number of health problems that twins had in old age. Specifically the twins were asked whether they had any of the following problems: abdominal or aortic aneurysm, brain aneurysm, rheumatoid arthritis, diverticulitis, emphysema, hemorrhoids or piles, hiatal hernia, kidney stone liver cirrhosis, prostate enlargement, cataracts, glaucoma, macular hole in retina, macular degeneration. For health behaviors, we used an indicator variable for having never smoked and an indicator variable equal to 1 if the twin reported that they drank beer daily, 3-6 times a week, or twice a week.

Statistical Analysis

To illustrate the within-MZ twins approach, suppose that the health measure of twin i in pair j (Y_{ij}) is related to schooling (S_{ij}), unobserved factors (such as ability, time preference, childhood family characteristics) that are common to both MZ twins in a given pair (μ_j) and an unobserved stochastic term (ε_{ij}).

$$Y_{ij} = \beta S_{ij} + \mu_j + \varepsilon_{ij} \quad (1)$$

An OLS regression of relation (1) provides an estimate of the association between schooling and health, which is a biased estimate of the causal impact for two reasons. First, β is likely to be biased because schooling is partially related to unobserved factors that affect both schooling and health directly (with the bias upward if the unobserved factors affect both schooling and health in the same direction). Second, classical measurement error in schooling causes β to be downward-biased. To control for measurement error, self-reported schooling can be instrumented with the co-twin's report (or some other report, as by the twins' children in Behrman et al. 1994) of the other twin's schooling. The unobserved factors influences can be controlled through the within-MZ twin estimator.

$$Y_{1j} - Y_{2j} = \beta(S_{1j} - S_{2j}) + (\varepsilon_{1j} - \varepsilon_{2j}) \quad (2)$$

The influence of unobserved factors μ_j is differenced out as MZ twins are genetically identical, the vast majority of MZ twins grow up together, and share many other socioeconomic contexts such as parental families, schools and neighborhoods.

There are two problems with within-MZ twins estimators in relation (2) that may still lead to biased estimates. First, the estimates are based on MZ twin pairs that differ in schooling attainment. If differences in schooling attainment are due to factors that also directly affect health, then the estimates will be biased (Bound and Solon 1999, Behrman et al. 2011, Kohler et al. 2011). Specifically β is upward- (downward-) biased if the factor leading to differences in schooling attainment affects schooling and health in the same (opposite) direction (see Behrman et al. 2011, Kohler et al. 2011).

For example, schooling differences between MZ twins may be due to differences in youth health, that also directly affect later life health. Both the MATR and MTR surveys asked twins whether they have ever suffered from any health conditions such as heart failure, high blood pressure, asthma, hearing impairment, depression, and if so at what age these conditions were diganosed. We attempted to control for early life health differences as a source of bias, by including an indicator equal to one if the twin was diagnosed with a health condition at age 16 or younger in all cross-sectional and within-MZ twins regressions. Second, the attenuation bias due to measurement error is exacerbated by first differencing. To deal with measurement error, we followed the Ashenfelter and Krueger (1994) strategy of instrumenting the difference in self-reported schooling with the difference in the co-twin's report of the other's schooling.

Relation (4) was used to estimate the effect of spousal schooling on own health. Relation (4) is the same as relation (1), but also includes the schooling of the spouse of twin i in pair j (Sp_{ij}).

$$Y_{ij} = \beta S_{ij} + \delta Sp_{ij} + \mu_j + \varepsilon_{ij} \qquad (4)$$

The cross-spouse effect δ is a biased estimate of the causal effect, because of the unobserved factors that are correlated with spousal schooling and also directly affect own health. For example, high-ability individuals tended to have better-schooled spouses and better own health. The bias due to unobserved factors is differenced out by contrasting the health of MZ twins with different spousal schooling in relation (5), and approach used for the first time to our knowledge in this and the concurrent study by Behrman et al. (2013).

$$Y_{1j} - Y_{2j} = \beta(S_{1j} - S_{2j}) + \delta(Sp_{1j} - Sp_{2j}) + (\varepsilon_{1j} - \varepsilon_{2j}) \qquad (5)$$

Because within-MZ schooling differences exist over most schooling levels, MZ within estimates are likely to be closer to average treatment effects (ATE) than local average treatment effects (LATE) (Moffitt 2009). IV-approaches that rely on variation in compulsory schooling as a first-stage instrument to predict schooling (Angrist and Krueger 1991; Lleras-Muney 2005), for example, yield local average treatment effects (LATE) relevant for individuals who are at the margin to be affected by the instruments used (e.g., the margin of completing only compulsory schooling levels) but not average treatment effects for broader populations beyond this margin (Behrman et al. 2011, Kohler et al. 2011, Lundborg 2013). ATE (or approximations thereof) that are provided by within-MZ approaches are likely preferable to LATE estimates that result from the use of such instruments as compulsory schooling regulations.

Results

Table 1 provides descriptive statistics. The average age of twins at the time of survey was 46 in the MTR, 52 in the MATR and 74 years in the NAS-NRC. The majority of twins in the MTR and MATR were female. The average grades of schooling was similar in the MTR and NAS-NRC datasets (approximately 14 grades), and lower in the MATR (13.52 grades). Only 2 percent of twins did not graduate from high school in the MTR, compared to 9 percent and 12 percent in the MATR and NAS-NRC. The proportion of twins that graduated from high school, had some post-high school schooling, and had bachelors degrees was similar across datasets. Although the MATR had the largest sample of twins, it had the least variation in grades of schooling. On average twin pairs had a difference of 0.69 grades and 71 percent of twin pairs had no difference in grades. In comparison in the MTR and NAS-NRC, 51 percent and 43 percent of twin pairs had the same grades of schooling, and on average the twin pairs had differences of 1.10 and 1.40 grades.

Average self-reported health was similar in the MATR and MTR (4.34 and 4.37) but substantially lower in the NAS-NRC (3.48). Average BMI was similar in all datasets. Approximately one half of twins were overweight in the MATR and NAS-NRC, whereas in the MTR only 38 percent of twins were overweight. In the MTR 16 percent of twins had an early health problem, more than double the proportion in the MATR.

There was also fairly substantial within-twin pair variation in the health measures. The absolute within-twin pair difference in self-reported health was 0.48 in the MTR, 0.58 in the MATR and and much higher in the NAS-NRC (0.84). In all three datasets, the absolute within-twin pair difference in BMI was over 2 units, and over 18 (8) percent of twin pairs were discordant on overweight (obesity) status. 20 percent of twin pairs differed on smoking status in the MATR and NAS-NRC. The absolute within-twin pair diffference in number of health problems for the elderly NAS-NRC twins was 1.22. For specific health conditions such as ever had a diabetes, prostrate cancer, heart attack, the proportion of twin pairs that differed was 11, 15 and 22 percent respectively.

The main results are in Table 2. All cross-sectional regressions control for age, age squared as well as gender and early health problems in the MATR and MTR. The estimates in column 1 indicate that schooling was associated with better health outcomes in the three data sets, in terms of higher self-reported health, lower BMI and lower probability of being overweight and obese. Moreover, in the elderly sample of NAS-NRC twins, more schooling was associated with having fewer health-limiting problems. However, there was no relationship between schooling and specific health conditions (heart attack, diabetes, high blood pressure, prostate cancer) and the total number of health problems in old age. Schooling was associated with better health behaviors. More-schooled individuals had higher probabilities of having never smoked, and of undertaking some exercise in leisure time. More schooling was associated with a lower probability of drinking beer frequently in the NAS-NRC. Surprisingly, however, more schooling was associated with more alcohol consumption in the MATR. The positive association may be plausible if people believed that there were health benefits of modest alcohol consumption or if alcohol was more affordable to more-schooled individuals due to higher income. These associations increased in magnitude, once measurement error was controlled (column 2) by instrumenting self-reported schooling with reports from co-twins.

Column 3 shows within-MZ-twins estimates, which also controlled for the early health problems indicator in the MATR and MTR. Compared to the cross-sectional estimates in column 1, the majority of the within-MZ twins estimates were substantially smaller in magnitude and statistically insignificant. This suggests that there was no causal effect of schooling. Rather the cross-sectional associations were cofounded by unobserved factors that affected schooling and health directly. There were however two exceptions. First, schooling was still significantly associated with better self-reported health and lower probabilities of being overweight in the MATR, even when controlling for unobserved factors in column 3. An extra grade of schooling increased self-reported health by 0.05 units, on a scale of 0-5, and lowered probabilities of being overweight by 1.2 percent. When self-reported schooling was instrumented using co-twins' reports of the others' schooling to control for random measurement error in column 4, the point estimates increased to 0.097 and 4.7 percent respectively, although the former estimate was no longer significant. Second, schoolings appeared to positively impact self-reported health in old age, though only about half as much as the cross-sectional association in column 1. An extra grade of schooling increased self-reported health by 0.035 units, on a scale of 0-5 for elderly men in the NAS-NRC sample. The point estimates for the other health measures also increased in absolute magnitudes, when instrumenting for measurement error in column 4. They all remained statistically insignificant, with the exception of exercise in the MATR, which was significant at the 10 percent level.

Appendix table 1 reports results by gender for MTR and MATR, and table 2 summarizes the difference in schooling coefficients for women and men. The cross-sectional analyses again indicated that more schooling was associated with better health outcomes and behaviors, with the possible exception of alcohol consumption. For self-reported health, the coefficient on the "early health problem indicator" was negative and significant, indicating that twins that suffered from a health condition in youth reported worse health compared to those who did not suffer from any health problems. There were only two significant gender differences in the cross-sectional schooling coefficient estimates: larger absolute magnitudes of associations of an extra grade of schooling with BMI and with never smoked for women than for men in the MATR sample. But neither of these gender differences was significant in the within-MZ estimates. On the other hand, every additional grade of schooling reduced probabilities of being overweight in the MTR sample by 5.6% more for men than for women. In the MATR, to investigate heterogeneous effects by age we interacted grades of schooling with indicator variables for being (1) 25-44, (2) 45-64 and (3) 65-75 years of age. There was no significant differential effect of schooling by these age groups (see appendix table 3).

As noted above, IV estimates reflect LATEs. For example, IV estimates based on changes in compulsory schooling laws estimate effects of additional grades of schooling at the low ends of schooling distributions. In contrast, because within-MZ twin schooling differences are across schooling distibutions, we estimated effects of both low and high levels of schooling with non-linear specifications for schooling as in table 3. (No measurement-error-corrected IV estimates are presented because with binary variables, measurement error is non-classical: individuals in the lowest educational category cannot under-report their education and individuals in the top categories cannot over-report their education.) Non-linearities were introduced through dichotomous variables indicating: (1) whether graduated from high school or lower, (2) whether had some post-high school schooling and (3) whether had bachelors degree or higher. The cross-sectional results indicate that those with some post-high school schooling and those with bacelors degrees had better health outcomes and behaviors compared to those who have graduated from high school or lower. The coefficient on having bachelors degrees or higher was almost twice the coefficient on some post-high school schooling, suggesting that there were larger health returns from completing college. Controlling for unobserved factors using within-MZ twins estimators (column 3) produced estimates that generally were smaller in magnitude and statistically insignificant. As before there was still some indication that schooling led to better self-reported health and lower probabilities of being overweight in the MATR. Individuals with bachelors degrees or higher had higher (lower) self-reported health (probabilities of being overweight) by 0.257 units (6.4 percent) compared to individuals who had graduated high school or lower. In the NAS-NRC sample, there was some evidence that the representation of schooling matters for the obtained results. Using linear grades of schooling (table 2), an extra grade of schooling led to higher reported health. However, using a non-linear representation, the within-MZ twin estimates for self-reported health were insgnificant. Moreover, the within-MZ twins estimates for BMI, suggested that men who had at least bachelors degrees had lower BMI of 0.574 units compared to men who graduated high school or lower. In comparison, there was no significant association of grades of schooling and BMI within-MZ twin pairs in table 2.

Table 4 reports estimates of spousal schooling on own health. All cross-sectional estimates in columns 1 and 2 controlled for age, age squared, gender and early health problems. The estimates in column 1 show that spousal schooling was associated with better self-reported health, lower BMI and lower probabilities of being overweight and obese. Spousal schooling also was associated with increased probabilities of having never smoked, undertaking some exercise in leisure time and alcohol consumption. The absolute magnitudes of estimated associations with spousal schooling fall in column 2, when own schooling was introduced. Spousal schooling still had positive associations with health outcomes, but no longer had any significant associations with health behaviors. For health outcomes in the MTR, it was surprising that the associations with spousal schooling were larger than those with own schooling. One would usually expect the opposite, which was found in the MATR. The within-MZ twins estimates in columns 3 and 4 were all insignificant, suggesting that the associations in columns 1 and 2 were due to influences of unobserved factors. Appendix table 4 shows estimates of interactions between the female indicator and (1) spousal and (2) own schooling in cross-sectional and within-MZ twins regressions. The cross-sectional regressions indicated that husband's schooling had greater effects in reducing BMI, probabilities of being overweight, obese in the MTR, and increasing probabilities of never smoking, engaging in frequent exercise and alcohol consumption compared to wives' schooling in the MATR. None of these differences were statistically significant when controlling for unobserved factors within-MZ twin pairs. In the MATR there were significant differential effects of spousal schooling on BMI and probabilities of being overweight, within-MZ twin pairs. An extra grade of wife's schooling reduced BMI by 0.155 units and the probability of being overweight by 1.8 percent compared to husband's schooling. In the MTR, wife's schooling increased self-reported health by 0.53 units compared to husband's schooling, within-MZ twin pairs.

Discussion

This study uses large US twins datasets and found that more schooling generally was associated with better health outcomes and health behaviors with the possible exception of more alcohol consumption. The within-MZ twins estimates indicated that more schooling did not causally improve health behaviors. Rather the positive cross-sectional relationships appeared to be reflecting influences of unobserved factors that affected schooling and health behaviors directly. This finding is in line with previous US twin-based evidence (Fujiwara and Kawachi 2009, Lundborg 2013), as well as studies of other countries (Behrman et al. 2011). We found, however, some evidence that more schooling causally led to some better health outcomes, though generally with smaller effects than suggested by cross-sectional associations. More schooling was still significantly associated with better self-reported health and lower probabilities of being overweight, once unobserved factors were controlled with within-MZ twin estimators. The beneficial effect was mainly driven by completing college. Like Lundborg (2013), we also found the functional form of schooling can lead to different results. Using linear grades of schooling, the within-MZ twins estimates showed no association between schooling and BMI for old-age men. In contrast, using a non-linear functional form, we found that old-age men with bachelor degrees had lower BMI compared to old-age men who had graduated from high school or lower. Finally, spousal schooling was associated with better health outcomes and behaviors, but the associations disappeared once unobserved cofounders were removed within-MZ twin pairs. We also did not find any consistent evidence for gender differences in effects of own schooling and spousal schooling.

This study has several limitations. First, within-MZ twins estimates may still be biased if differences in schooling attainment were due to factors that also directly affected health. We attempted to control for one possible factor- early life health differences. Differences in schooling attainment may be due to other factors such as differences in parental treatment or school quality, that also have direct affects on health, which we cannot rule out. But if these factors were positively associated with both schooling and health, the fact that we cannot control for them led our within-MZ estimates to be upper bounds on absolute magnitudes of true schooling effects. Because our results in many cases indicated no or only weak effects of schooling on health, our basic conclusion that the causal effects of schooling on health are generally weak or non-existent is not affected by such biases. Second, we used data on twins born in specific states and/or from certain birth cohorts, who were therefore unlikely to have been representative of the general US population. To assess the representativeness of the twins data, we compared the MATR (MTR) twins to nationally-representative samples of white individuals aged 25-75 (39-58) in the 1987, (1994) Behavioral Risk Factor Surveillance System (BRFSS). The elderly male NAS-NRC twins were compared to elderly male veterans in the 2000 BFRSS. Summary statistics from the BFRSS in appendix table 5 show that the twin samples tended to have higher schooling attainment than the nationally-representative samples from these birth cohorts. The MATR and MTR twins data sets also over-represented women. Average self-reported health was lower in the 1994 and 2000 BFRSS compared to the MTR and NAS-NRC twins. BMI was similar in the BFRSS and twins data sets. The proportion of individuals who had never smoked was slightly lower in the 1987 and 2000 BFRSS compared to the MATR and NAS-NRC twins. Appendix table 6 presents cross-sectional schooling estimates, where we combined the twins and BFRSS data and included an interaction between grades of schooling and a dummy variable for the twins data to test whether schooling coefficients differed in the twins and BFRSS data sets. In the combined MATR-BFRSS data set, there was a significant difference in schooling coefficients on BMI, probabilities of being overweight, obese for twins and individuals in the BFRSS. There was no significant difference for probabilities of having never smoked. In the combined MTR-BFRSS data set, for BMI and the incidence of overweight and obesity, the interaction was insignificant, but significant for self-reported health. In the combined NAS-NRC BFRSS dataset, the interaction was significant only for probabilities of never smoking. In general, although the results may not be fully generalizable, differences between the twins and BFRSS samples did not cause biases in our within-MZ twins estimates if these differences from representative samples were due to the unobserved factors that were controlled for in the within-MZ estimates. Third, we did not investigate possible mechanisms under which schooling improved health except that if we included income in our estimates the basic results for schooling were not altered much (estimates available from authors), suggesting that schooling was not working primarily through income.

Despite these limitations, this study provides more evidence on causal effect of schooling on health, to a literature where studies for the US and other developed countries either found no or a small causal effect of schooling. Our first finding that schooling improved somewhat self-reported health, though not as much as indicated by cross-sectional associations, is consistent with Lundborg (2013) but contrasts with instrumental variable estimates for the US (Mazumder 2008), UK (Clark and Royer 2013), and Denmark (Arendt 2005), where no effect was found. Our second finding that more schooling lowered probabilities of being overweight, differed from within-MZ twins estimates for the US where no causal relationship was found (Fujiwara and Kawachi 2009, Lundborg 2013), and for Australia where schooling reduced probabilities of being overweight for men but not for women (Webbink et al. 2010). Instrumental variables estimates for the UK (Clark and Royer 2013), Denmark (Arendt 2005) also showed no evidence of any causal relationship. However, our finding is consistent with instrumental variable estimates for the US (Grabner 2008, MacInnis 2006) and Germany (Kemptner et al. 2011).

References

Anderson, L. S., Beverly, W. T., Corey, L. A., & Murrelle, L. (2002). The Mid-Atlantic Twin Registry. *Twin Research, 5(5),* 449–455.

Angrist, J. D. & Krueger, A. B. (1991). Does Compulsory School Attendance Affect Schooling and Earnings. *Quarterly Journal of Economics*, *106(4),* 979-1014.

Arendt, J. N. (2005). Does education cause better health? A panel data analysis using school reforms for identification. *Economics of Education Review*, *24*(2), 149-160.

Ashenfelter, O., & Krueger, A. (1994). Estimates of the economic return to schooling from a new sample of twins. *American Economic Review, 84*(5), 1157–1173.

Behrman, J.R., Kohler, H-P., Jensen, V.M., Pedersen, D., Petersen, I., Bingley, P., & Christensen, K. (2011). Does schooling reduce hospitalization and delay mortality? New evidence based on Danish twins. *Demography*. *48(4)*, 1347-1375.

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(6),* 1131-74.

Behrman, J.R., Xiong, Y., & Zhang, J. (2013). Schooling-Health Associations Misrepresent the Causal Effects of Schooling on Health: Evidence from Chinese Twins. Hong Kong: Chinese University of Hong Kong.

Bound, J., & Solon, G. (1999). Double trouble: on the value of twins-based estimation of the returns to schooling. *Economics of Education Review, 18(2)*, 169-182.

Buckles, K., Ofer, M., Morril, M.S. & Wozniak, A. (2012). The effect of college education on health. IZA Discussion Paper Number 6659.

Clark, D., & Royer, H. (2013). The Effect of Education on Adult Mortality and Health: Evidence from Britain. *American Economic Review, 103(6),* 2087-2120.

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

De Walque, D. (2007). Does education affect smoking behaviors?: Evidence using the Vietnam draft as an instrument for college education. *Journal of Health Economics, 26(5),* 877-895.

Grimard, F., & Parent, D.(2007). Education and Smoking: Were Vietnam War Draft Avoiders Also More Likely to Avoid Smoking? *Journal of Health Economics*, *26(5),* 896-926.

Grossman, M. (1972). On the concept of health capital and demand for health. *Journal of Political Economy*, *80(2),* 223-255.

Grabner, M. (2008). The Causal Effect of Education on Obesity. Evidence from Compulsory Schooling Laws. Mimeo, University of California at Davis.

Glymour, M. M, Kawachi, I., Jencks, C. S. & Berkman, L. F. (2008). Does childhood schooling affect old age memory or mental status? Using state schooling laws as natural experiments. *Journal of Epidemiology Community Health, 62(6),* 532–537.

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(13)*, 1310-1322.

Lleras-Muney, A. (2005). The Relationship between Education and Adult Mortality in the United States. *Review of Economic Studies*, *72(1)*, 189-221.

Lundborg, P. (2013) "The health returns to education- What can we learn from twins?" *Journal of Population Economics, 26(2),* 673-701.

Lykken, D., Bouchard, T., McGue, M. & A., T. (1990). The minnesota twin family registry – some initial findings. *Acta Geneticae Medicae et Gemellologiae, 39(1),* 35–70.

Kemptner D., Jurges, H. & Reinhold, S. (2011). Changes in compulsory schooling and the causal effect of education on health: Evidence from Germany. *Journal of Health Economics, 30(2),* 340-354

Kohler, H.-P.; Behrman, J. & Schnittker, J. (2011). Social Science Methods for Twins Data: Integrating Causality, Endowments and Heritability. *Journal of Biodemography and Social Biology*, *57(1),* 88-141.

MacInnis, B. (2006). The Long-Term Effects of College Education on Morbidities: New Evidence from the Pre-Lottery Vietnam Draft. *Mimeo*, University of Michigan.

Mazumder, B. 2008. Does education improve health? A reexamination of the evidence from compulsory schooling laws. *Economic Perspectives, 33(2)*, 2–16.

Moffitt, R. A. (2009). Issues in the Estimation of Causal Effects in Population Research, With an Application to the Effects of Teenage Childbearing. In Engelhardt, H., Kohler, H-P & Fürnkranz-Prskawetz, A. (eds.) *Causal Analysis in Population Studies*, Berlin: Springer Verlag , 9-29

Page, W.F. (2002). The NAS-NRC Twin Registry of WWII Military Veteran Twins. Twin *Research, 5(5),* 493-496

Webbink, D., Martin, N. & Visscher, P. (2010). Does Education reduce the probability of being overweight? *Journal of Health Economics*, *29(1)*, 29-38.

Table 1: Descriptive Statistics

	MATR	MTR	NAS-NRC

	(1)	(2)	(3)
Background Characteristics			
Age	52.41 (15.04)	46.63 (5.49)	74.16 (2.76)
25-34	0.22 (0.40)		
35-44	0.12 (0.33)	0.42 (0.49)	
45-54	0.12 (0.33)	0.47 (0.50)	
55-64	0.29 (0.45)	0.11 (0.31)	
65-75	0.26 (0.44)		0.66 (0.47)
76-85			0.30 (0.46)
Female	0.71 (0.45)	0.65 (0.47)	0.00 (0.00)
Grades of schooling	13.52 (2.10)	14.03	14.02 (3.00)
Absolute within-twin difference in grades of schooling	0.69 (1.17)	1.10 (1.52)	1.40 (0.87)
0 grade difference	0.71 (0.46)	0.51 (0.50)	0.43 (0.87)
1 grade difference		0.19 (0.39)	0.23 (0.42)
2 grades difference	0.25 (0.43)	0.14 (0.35)	0.14 (0.35)
3 grades difference	0.00 (0.05)	0.05 (0.22)	0.06 (0.24)
4 grades difference	0.04 (0.18)	0.07 (0.25)	0.07 (0.26)
5 grades difference or more	0.01 (0.07)	0.04 (0.18)	0.06 (0.24)
Less than 12 grades	0.09 (0.28)	0.02 (0.12)	0.12 (0.32)
12 grades	0.34 (0.48)	0.33 (0.47)	0.25 (0.43)
13-15 grades	0.26 (0.44)	0.30 (0.46)	0.25 (0.43)
16 grades or more	0.31 (0.46)	0.35 (0.48)	0.38 (0.48)
Co-twin's grades of schooling	13.43 (2.12)	13.91 (2.37)	
Spouse's grades of schooling	13.57 (2.29)	13.81 (2.47)	
Early health problem	0.07 (0.25)	0.16 (0.36)	
Health Outcomes			
Self-reported health	4.34 (0.86)	4.37 (0.67)	3.48 (0.98)
BMI	24.54 (4.27)	25.74 (4.49)	25.79 (3.31)

Table 1 Continued

Overweight	0.38 (0.47)	0.50 (0.50)	0.58 (0.49)
Obese	0.10 (0.30)	0.16 (0.36)	0.09 (0.29)
Limiting activities			14.54 (4.90)
Ever had heart attack			0.16 (0.37)
Ever had stroke			0.09 (0.29)
Ever had high blood pressure			0.46 (0.50)
Ever had diabetes			0.15 (0.36)
Ever had prostrate cancer			0.12 (0.33)
Number of health problems			1.97 (1.48)
Health Behaviors			
Never Smoked	0.53 (0.50)		0.37 (0.48)
Exercise	0.52 (0.50)		
Alcohol Consumption	2.48 (5.81)		0.36 (0.47)
Absolute within-twin pair difference in			
Self-reported health	0.58 (0.76)	0.48 (0.63)	0.82 (0.78)
BMI	2.08 (2.28)	2.49 (2.44)	2.28 (2.01)
Number of health problems			1.22 (1.05)
Limiting activities			3.84 (4.29)
Alcohol consumption	2.45 (5.68)		
Proportion of twin pairs that differ in			
Overweight	0.19 (0.40)	0.23 (0.42)	0.30 (0.45)
Obese	0.08 (0.28)	0.14 (0.34)	0.11 (0.31)
Ever had heart attack			0.22 (0.41)
Ever had stroke			0.13 (0.33)
Ever had high blood pressure			0.26 (0.44)
Ever had diabetes			0.11 (0.31)
Ever had prostate cancer			0.15 (0.35)
Never Smoked	0.20 (0.40)		0.20 (0.40)
Exercise	0.33 (0.47)		
Alcohol consumption			0.33 (0.47)

Notes: Standard deviations in parentheses.

Table 2: Cross-Sectional and Within-MZ Twins Estimates

	Cross-Section		Within-MZ Twins	
	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)
Self-reported health				
MATR	0.075	0.081	0.05	0.097
	(.007)***	(.009)***	(.017)**	(0.064)
	[4238]	[4238]	[2164]	[2164]
MTR	0.045	0.059	0.014	0.018
	(0.008)***	(0.012)***	(0.014)	(0.022)
	[1310]	[1310]	[655]	[655]
NAS-NRC	0.067		0.035	
	(0.008)***		(0.018)*	
	[1868]		[934]	
BMI				
MATR	-0.243	-0.286	-0.063	-0.145
	(.037)***	(.045)***	(0.060)	(0.215)
	[4272]	[4272]	[2136]	[2136]
MTR	-0.182	-0.248	0.061	0.035
	(0.061)***	(0.091)**	(0.075)	(0.113)
	[1290]	[1290]	[645]	[645]
NAS-NRC	-0.115		-0.056	
	(0.030)***		(0.047)	
	[1844]		[922]	
Overweight				
MATR	-0.027	-0.03	-0.012	-0.047
	(.004)***	(.005)***	(.006)*	(.026)*
	[4272]	[4272]	[2136]	[2136]
MTR	-0.014	-0.023	0.011	0.011
	(0.007)**	(0.010)**	(0.010)	(0.019)
	[1290]	[1290]	[645]	[645]
NAS-NRC	-0.014		-0.005	
	(.004)***		(0.008)	
	[1844]		[922]	
Obese				
MATR	-0.013	-0.016	-0.002	-0.004
	(.002)***	(.003)***	-0.005	-0.017
	[4272]	[4272]	[2136]	[2136]
MTR	-0.009	-0.012	-0.001	-0.002
	(0.005)*	(0.007)*	(0.007)	(0.011)
	[1290]	[1290]	[645]	[645]

Table 2 Continued

NAS-NRC	-0.007		-0.002	
	(.002)**		(0.004)	
	[1844]		[922]	
Ever had heart attack				
NAS-NRC	-0.003		0.001	
	(0.003)		(0.006)	
	[1882]		[922]	
Ever had diabetes				
NAS-NRC	-0.005		0.007	
	(0.003)		(0.005)	
	[1876]		[938]	
Ever had high blood pressure				
NAS-NRC	-0.007		-0.004	
	(0.004)		(0.007)	
	[1882]		[941]	
Ever had prostrate cancer				
NAS-NRC	-0.006		-0.004	
	(0.015)		(0.007)	
	[1342]		[671]	
Limiting activities				
NAS-NRC	-0.262		-0.029	
	(0.042)***		(0.083)	
	[1726]		[863]	
Number of health problems				
NAS-NRC	-0.007		-0.031	
	(0.015)		(0.026)	
	[1342]		[671]	
Never Smoked				
MATR	0.018	0.022	0.022	0.013
	(.005)***	(.005)***	(0.007)	(0.027)
	[4126]	[4126]	[2063]	[2063]
NAS-NRC	0.031		0.007	
	(0.004)***		(0.007)	
	[1902]		[951]	
Exercise				
MATR	0.037	0.041	0.01	0.076
	(.004)***	(.005)***	(0.010)	(.036)*
	[4240]	[4240]	[2120]	[2120]

Table 2 Continued

Alcohol Consumption				
MATR	0.091	0.116	-0.058	-0.347
	(.054)*	(0.064)*	(0.120)	(0.534)
	[3392]	[3392]	[1696]	[1696]
NAS-NRC	(0.020)		0.001	
	(0.005)***		-0.01	
	[1048]		[524]	

Notes: Cross sectional regressions in the MATR and MTR control for age, age squared, gender and early health problems. Cross-sectional regressions for NAS-NRC control for age and age squared. For cross-sectional IV estimates, twin 1's schooling is instrumented by twin 2's report of twin 1's schooling and vice versa. All Within-MZ twins estimates for MATR and MTR control for early health problems. For within-MZ twins IV estimates, the difference in self-reported schooling is instrumented by the difference in the co-twin's report of the other's schooling. Standard errors clustered by twin pairs in (.). ***significant at 1% ** significant at 5%; * significant at 10%. Sample size in [.]

Table 3: Cross-Sectional and Within-MZ Twins Estimates, Non-Linear Schooling

	Cross-Section OLS	Within-MZ Twins OLS
Self-reported health		
MATR		
Some college	0.174	0.101
	(0.036)***	(0.063)
Degree	0.305	0.257
	(0.033)***	(0.081)***
	[4328]	[2164]
MTR		
Some college	0.124	-0.021
	(0.047)***	(0.065)
Degree	0.262	0.061
	(0.047)***	-0.078
	[1316]	[655]
NAS-NRC		
Some college	0.225	0.055
	(0.060)***	(0.093)
Degree	0.417	0.158
	(0.057)***	(0.123)
	[1868]	[934]
BMI		
MATR		
Some college	0.373	-0.232
	(0.193)**	(0.203)
Degree	-1.06	-0.251
	(0.179)***	(0.320)
	[4272]	[2136]
MTR		
Some college	-0.356	0.132
	(0.375)	(0.337)
Degree	-0.918	0.415
	(0.358)**	(0.406)
	[1290]	[645]
NAS-NRC		
Some college	-0.557	-0.373
	(0.215)**	(0.243)
Degree	-0.727	-0.574
	(0.210)***	(0.327)*
	[1844]	[922]

Table 3 Continued

Overweight		
MATR		
Some College	-0.03	-0.032
	(0.021)	(0.024)
Degree	-0.123	-0.064
	(0.020)**	(0.032)**
	[4272]	[2136]
MTR		
Some college	-0.031	0.049
	(0.037)	(0.041)
Degree	-0.062	0.114
	(0.037)*	(0.050)**
	[1290]	[645]
NAS-NRC		
Some college	-0.048	-0.03
	(0.031)	(0.044)
Degree	-0.069	-0.043
	(0.029)**	-0.055
	[1844]	[922]
Obese		
MATR		
Some College	-0.03	-0.016
	(0.013)**	(0.020)
Degree	-0.057	-0.021
	(0.011)***	-0.025
	[4272]	[2136]
MTR		
Some college	-0.028	-0.059
	(0.045)	(0.059)
Degree	-0.039	-0.015
	(0.023)*	-0.041
	[1290]	[645]
NAS-NRC		
Some college	-0.055	-0.043
	(0.018)***	(0.024)*
Degree	-0.053	-0.045
	(0.018)***	(0.031)
	[1844]	[922]

Table 3 Continued

Limiting Activities		
NAS-NRC		
Some college	-0.69	-0.1
	(0.315)***	(0.469)
Degree	-1.58	0.007
	(0.294)***	(0.607)
	[1726]	[863]
Never Smoked		
MATR		
Some college	0.015	0.019
	(0.021)	(0.027)
Degree	0.097	0.011
	(0.022)***	-0.033
	[4126]	[2063]
NAS-NRC		
Some college	0.041	-0.021
	(0.029)	(0.034)
Degree	0.219	0.053
	(0.029)***	(0.045)
	[1902]	[951]
Exercise		
MATR		
Some college	0.087	0.002
	(0.021)***	(0.037)
Degree	0.178	0.067
	(0.020)***	(0.046)
	[4240]	[2120]
Alcohol consumption		
MATR		
Some college	0.391	-0.656
	(0.282)	(0.438)
Degree	0.303	-0.142
	(0.277)	(0.617)
	[3392]	[1696]
NAS-NRC		
Some college	-0.059	0.045
	(0.043)	(0.062)
Degree	-0.147	-0.063
	(0.038)***	(0.084)
	[1048]	[524]

Notes: The omitted schooling category is high school graduate or lower. Cross sectional regressions in the MATR and MTR control for age, age squared, gender and early health problems. Cross-sectional regressions for NAS-NRC control for age and age squared. Within-MZ twins estimates for MATR and MTR control for early health problems.

Standard errors clustered by twin pairs in (.). ***significant at 1% ** significant at 5%; * significant at 10%. Sample size in [.]

Table 4: Cross-Sectional and Within-MZ Twins Estimates of Spousal Schooling

	Cross-Section		Within-MZ Twins	
	OLS	OLS	OLS	OLS
	(1)	(2)	(3)	(4)
Self-reported health				
MATR				
Spousal schooling	0.046	0.022	0.014	0.011
	(0.010)***	(0.011)**	(0.012)	(0.012)
Own schooling		0.049		0.025
		(0.012)***		(0.024)
N	1674	1674	837	837
MTR				
Spousal schooling	0.045	0.035	0.009	0.007
	(0.009)***	(0.010)***	-0.016	-0.016
Own education		0.023		0.019
		(0.011)**		(0.019)
N	868	868	434	434
BMI				
MATR				
Spousal schooling	-0.286	-0.197	-0.024	-0.005
	(0.050)***	(0.052)***	(0.045)	(0.044)
Own schooling		-0.182		-0.174
		(0.064)***		(0.077)**
N	1644	1644	822	822
MTR				
Spousal schooling	-0.256	-0.207	0.079	-0.075
	(0.064)***	(0.067)***	(0.058)	(0.060)
Own schooling		-0.108		-0.33
		(0.073)		(0.990)
N	846	846	434	434
Overweight				
MATR				
Spousal schooling	-0.03	-0.021	-0.005	-0.004
	(0.005)***	(0.006)***	(0.006)	(0.006)
Own schooling		-0.019		-0.006
		(0.007)**		(0.011)
N	1644	1644	822	822

Table 4 Continued

MTR				
Spousal schooling	-0.021	-0.018	0.002	0.001
	(0.007)***	(0.008)**	(0.008)	(0.008)
Own schooling		-0.005		0.008
		(0.009)		(0.013)
N	846	846	434	434
Obese				
MATR				
Spousal schooling	-0.015	-0.011	0.000	0.001
	(0.004)***	(0.004)***	(0.005)	(0.005)
Own schooling		-0.009		-0.009
		(0.005)*		(0.007)
N	1644	1644	822	822
MTR				
Spousal schooling	-0.012	-0.008	-0.002	-0.001
	(0.005)**	(0.005)	(0.007)	(0.007)
Own schooling		-0.009		-0.005
		(0.006)		(0.010)
N	846	846	434	434
Spousal schooling	0.012	0.004	-0.003	-0.001
	(0.006)***	(0.007)	(0.006)	(0.006)
Own schooling		0.016		-0.014
		(0.008)**		(0.011)
N	1648	1648	824	824
Exercise				
Spousal schooling	0.017	0.003	-0.006	0
	(0.006)***	(0.006)	(0.009)	(0.009)
Own schooling		0.029		-0.001
		(0.007)***		(0.017)
N	1628	1628	814	814
Alcohol Consumption				
Spousal schooling	0.134	0.083	0.128	0.157
	(0.056)**	(0.067)	(0.116)	(0.124)
Own schooling		0.109		-0.256
		(0.084)		(0.171)
N	1338	1338	669	669

Notes: Cross sectional regressions control for age, age squared, gender and early health problems. Within-MZ Twins regressions control for early health problems. Standard errors clustered by twin pairs in (.). ***significant at 1% ** significant at 5%; * significant at 10%.

Appendix Table 1: Cross-Sectional and Within-MZ Twins Estimates, By Gender

	Men				Women			
	Cross-Section		Within-MZ Twins		Cross-Section		Within-MZ Twins	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Self-reported health								
MTR: Schooling	0.041	0.041	0.033	0.031	0.049***	0.075	-0.004	0.007
	(0.012)***	(0.018)**	(0.020)	(0.029)	(0.011)***	(0.017)***	(0.020)	(0.032)
Early health Problem	-0.111	-0.111	0.002	0.004	-0.152	-0.161	-0.12	-0.13
	(0.082)	(0.082)	(0.113)	(0.110)	(0.071)***	(0.070)***	(0.094)	(0.094)
N	464	464	232	232	846	846	423	423
MATR: Schooling	0.076	0.081	0.021	0.128	0.076	0.084	0.062	0.05
	(0.011)***	(0.015)***	(0.036)	(0.076)*	(0.009)***	(0.011)***	(0.021)***	(0.097)
Early health Problem	-0.248	-0.252	-0.265	-0.265	-0.296	-0.299	-0.092	-0.089
	(.118)***	(.111)**	(0.151)*	(0.147)	(0.067)***	(0.067)***	(0.089)	(0.089)
N	1262	1262	631	631	3066	3066	1533	1533
BMI								
MTR: Schooling	-0.165	-0.225	-0.04	0.031	-0.201	-0.272	0.148	0.038
	(0.075)**	(0.114)*	(0.108)	(0.134)	(0.092)**	(0.135)**	(0.103)	(0.183)
Early health Problem	-0.848	-0.865	-0.048	-0.04	-0.454	-0.428	-0.34	-0.323
	(0.421)**	(0.421)**	(0.405)	(0.401)	(0.438)	(0.442)	(0.333)	(0.335)
N	462	462	231	231	828	828	414	414
MATR: Schooling	-0.091	-0.123	0.071	0.187	-0.304	-0.353	-0.121	-0.379
	(0.056)*	(0.067)**	(0.143)	(0.342)	(0.048)***	(0.057)***	(0.059)**	(0.280)
Early health Problem	0.357	0.376	-0.157	-0.089	0.585	0.608	0.223	0.242
	(0.445)	(0.444)	(0.353)	(0.354)	(0.376)	(0.376)	(0.272)	(0.273)
N	1256	1256	628	628	3016	3016	1508	1508
Overweight								
MTR: Schooling	-0.013	-0.013	-0.019	-0.001	-0.015	-0.031	0.037	0.022
	(0.010)	(0.014)	(0.016)	(0.021)	(0.010)*	(0.013)**	(0.011)***	(0.017)
Early health Problem	-0.15	-0.145	-0.082	-0.081	-0.038	-0.033	-0.05	-0.047
	(0.063)**	(0.063)	(0.068)	(0.069)	(0.048)	(0.048)	(0.050)	(0.046)
N	462	462	231	231	828	828	414	414
MATR: Schooling	-0.022	-0.026	-0.004	-0.01	-0.028	-0.031	-0.015	-0.058
	(0.008)***	(0.009)***	-0.014	-0.037	(0.005)***	(0.006)***	(0.007)**	-0.041
Early health Problem	0.006	0.009	-0.12	-0.12	0.04	0.041	-0.005	-0.002
	(0.063)	(0.063)	(0.072)*	(0.072)	(0.033)	(0.033)	(0.032)	(0.032)
N	1256	1256	628	628	3016	3016	1508	1508

Appendix Table 1 Continued

Obese								
MTR: Schooling	-0.009	-0.012	-0.006	-0.001	-0.01	-0.012	0.003	-0.003
	(0.008)	(0.011)	(0.010)	(0.015)	(0.006)	(0.009)	(0.009)	(0.015)
Early health Problem	-0.065	-0.064	-0.041	-0.04	-0.029	-0.029	-0.033	-0.032
	(0.036)*	(0.036)*	(0.039)	(0.039)	(0.035)	(0.035)	(0.036)	(0.036)
N	462	462	231	231	828	828	414	414
MATR: Schooling	-0.006	-0.007	-0.005	-0.012	-0.016	-0.02	-0.001	0.003
	(0.004)*	(0.004)	(0.008)	(0.021)	(0.003)***	(0.004)***	-0.006	-0.028
Early health Problem	-0.011	-0.011	0.019	-0.19	0.049	0.051	0.047	-0.046
	(0.027)	(0.027)	(0.019)	(0.019)	(0.024)**	(0.024)**	(0.026)*	(0.026)*
N	1256	1256	628	628	3016	3016	1508	1508
Never Smoked								
MATR: Schooling	0.047	0.057	-0.002	0	0.009	0.01	0.005	0.025
	(0.007)**	(0.009)**	(0.013)	(0.031)	(0.006)	(0.007)	(0.008)	(0.043)
Early health Problem	0.024	0.016	-0.02	-0.02	-0.074	-0.074	0.046	0.044
	(0.061)	(0.061)	(0.054)	(0.060)	(0.036)**	(0.037)**	(0.040)	(0.037)
N	1188	1188	594	594	2938	2938	1469	1469
Exercise								
MATR: Schooling	0.045***	0.049***	0.029*	0.072	0.035	0.038	0.002	0.095
	(0.006)	(0.008)	(0.017)	(0.045)	(0.005)***	(0.006)***	(0.012)	(0.059)
Early health Problem	0.031	0.028	0.1	0.092	-0.036	-0.037	-0.013	-0.017
	(0.059)	(0.059)	(0.090)	(0.090)	(0.036)	(0.036)	(0.047)	(0.048)
N	1250	1250	625	625	2990	2990	1495	1495
Alcohol Consumption								
MATR: Schooling	0.035	0.029	-0.07	0.045	0.13	0.174	-0.054	-0.748
	(0.104)	(0.132)	(0.239)	(0.690)	(0.062)**	(0.072)**	(0.138)	(0.872)
Early health Problem	-1.26	-1.26	0.208	0.198	0.296	0.275	0.281	0.318
	(0.743)*	(0.741)*	(0.651)	(0.649)	(0.433)	(0.433)	(0.403)	(0.406)
N	930	930	465	465	2462	2462	1231	1231

Notes: Cross sectional regressions control for age, age squared, and an indicator for early health problems. For cross-sectional IV estimates, twin 1's schooling is instrumented by twin 2's report of twin 1's schooling and vice versa. Within-MZ twins regressions control for early health problems. For within-MZ twins IV estimates, the difference in self-reported schooling is instrumented by the difference in the co-twin's report of the other's schooling. Standard errors clustered by twin pairs in (.). ***significant at 1% ** significant at 5%; * significant at 10%. Sample size in [.]

Appendix Table 2: Differential Effect of Schooling, Women versus Men

	Cross-Sectional OLS	Cross-Sectional IV	Within-MZ Twins OLS	Within-MZ Twins IV
Self-reported health				
MTR	0.008 (0.016)	0.034 (0.024)	-0.037 (0.028)	-0.024 (0.043)
MATR	0.00 (0.014)	0.003 (0.018)	0.041 (0.037)	-0.078 (0.123)
BMI				
MTR	-0.036 (0.119)	-0.047 (0.176)	0.188 (0.149)	0.007 (0.226)
MATR	-0.213 (0.074)***	-0.023 (0.088)***	-0.192 (0.155)	-0.566 (0.442)
Overweight				
MTR	-0.002 (0.014)	-0.018 (0.019)	0.056 (0.019)***	0.023 (0.027)
MATR	-0.006 (0.009)	-0.005 (0.010)	-0.011 (0.016)	-0.048 (0.055)
Obese				
MTR	-0.001 (0.009)	0.00 (0.014)	0.008 (0.014)	-0.002 (0.015)
MATR	-0.010 (0.005)**	-0.013 (0.006)***	0.004 (0.010)	0.015 (0.035)
Never smoked				
MATR	-0.038 (0.009)***	-0.047 (0.011)***	0.007 (0.015)	0.025 (0.053)
Exercise				
MATR	-0.010 (0.008)	-0.011 (0.010)	-0.027 (0.021)	0.023 (0.074)
Alcohol consumption				
MATR	0.095 (0.121)	0.145 (0.150)	0.016 (0.276)	-0.793 (1.11)

Notes: Standard errors in parentheses. ***significant at 1% ** significant at 5%; * significant at 10%.

Appendix Table 3: Heterogeneous Schooling Effects By Age

	Cross-Section	Within-MZ Twins
	OLS	**OLS**
Self-reported health		
Years of education	0.067	0.053
	(0.011)***	(0.029)*
(age 45-64)*education	0.023	0.001
	(0.016)	(0.043)
(age 65-75)*education	-0.002	-0.008
	(0.017)	(0.038)
N	4238	4238
BMI		
Years of education	-0.224	-0.092
	(0.069)***	(0.077)
(age 45-64)*education	-0.006	0.028
	(0.093)	(0.146)
(age 65-75)*education	-0.032	0.059
	(0.094)	(0.115)
N	4272	2136
Overweight		
Years of education	-0.022	-0.018
	(0.007)***	(0.011)*
(age 45-64)*education	-0.006	-0.006
	(0.009)	(0.014)
(age 65-75)*education	-0.005	0.031
	(0.010)	(0.017)*
N	4272	2136
Obese		
Years of education	-0.009	-0.007
	(0.004)**	-0.006
(age 45-64)*education	-0.005	-0.012
	(0.006)	(0.010)
(age 65-75)*education	-0.006	-0.014
	(0.007)	(0.013)
N	4272	2136

Appendix Table 3 Continued

Smoking		
Years of education	0.061	0.002
	(0.007)***	(0.014)
(age 45-64)*education	-0.052	-0.004
	(0.010)***	(0.017)
(age 65-75)*education	-0.072	0.007
	(0.011)***	(0.019)
N	4126	2063
Alcohol Consumption		
Years of education	0.077	0.114
	(0.090)	(0.175)
(age 45-64)*education	-0.077	-0.303
	(0.134)	(0.292)
(age 65-75)*education	0.139	-0.182
	(0.120)	(0.257)
N	3392	1696
Exercise		
Years of education	0.042	0.004
	(0.007)***	-0.017
(age 45-64)*education	-0.086	0.009
	(0.068)	(0.023)
(age 65-75)*education	-0.009	0.012
	(0.010)	(0.027)
N	4240	2120

Notes: Estimates are based on the regression: $Y_{ij}=b_o + b_1$ Own Schooling$_{ij}$ + b_2(Age45to64)$_{ij}$+ b_4(Age65-75)$_{ij}$ b_4Female$_{ij}$ + b_6EarlyHealthProblem$_i$ + b_6 (Own Schooling)$_{ij}$*(Age45to64)$_{ij}$ + b_7 ((Own Schooling$_i$)$_{ij}$*(Age65to75)$_{ij}$ + u_{ij}.. The omitted age group is 24-44. Standard errors clustered by twin pairs in parentheses. ***significant at 1% **significant at 5% *significant at 10%

Appendix Table 4: Differential Effect of Spousal and Own Schooling

	Cross-Section	Within-MZ Twins
Self-reported health		
MATR		
Spousal schooling	0.03 (0.23)	0.041 (0.028)
Own Schooling	-0.035 (0.025)	0.017 (0.054)
MTR		
Spousal schooling	-0.007 (0.021)	-0.053 (0.032)*
Own Schooling	0.017 (0.022)	-0.026 (0.039)*
BMI		
MATR		
Spousal schooling	-0.167(0.105)	-0.155 (0.085)*
Own Schooling	-0.047 (0.123)	-0.017 (0.157)
MTR		
Spousal schooling	-0.234 (0.136)*	-0.005 (0.111)
Own Schooling	-0.035 (0.149)	-0.009 (0.201)
Overweight		
MATR		
Spousal schooling	0.008 (0.015)	-0.008 (0.014)
Own Schooling	-0.005 (0.015)	-0.024 (0.025)
MTR		
Spousal schooling	-0.015 (0.017)	0.000 (0.017)
Own Schooling	-0.005 (0.018)	0.053 (0.026)**
Obese		
MATR		
Spousal schooling	-0.020 (0.008)**	-0.018 (0.011)
Own Schooling	-0.002 (0.009)	0.009 (0.016)
MTR		
Spousal schooling	-0.024 (0.011)**	-0.016 (0.014)
Own Schooling	0.001 (0.011)	0.007 (0.020)
Never Smoke		
MATR		
Spousal schooling	-0.028 (0.014)**	-0.013 (0.014)
Own Schooling	-0.008 (0.015)	-0.005 (0.023)
Exercise		
MATR		
Spousal schooling	0.035 (0.015)***	0.016 (0.014)
Own Schooling	-0.046 (0.015)***	-0.054 (0.038)

Appendix Table 4 Continued

Alcohol Consumption		
MATR		
Spousal schooling	0.507 (0.195)**	-0.017 (0.024)
Own Schooling	-0.232 (0.199)	-0.010 (0.029)

Notes: Estimates are based on the regression: $Y_{ij}=b_o + b_1$ Own Schooling$_{ij}$ + b_2(Spousal Schooling$_{ij}$ + b_3Female$_{ij}$ + $b_4$age$_{ij}$ + $b_5$age$^2_{ij}$ + b_6EarlyHealthProblem$_{ij}$ + b_7 (Own Schooling)$_{ij}$*(Female)$_{ij}$ + b_8 ((Spousal Schooling$_i$)$_{ij}$*(Female)j$_i$ + b_9 (Age)$_{ij}$*(Female)$_{ij}$ + b_{10} (age^2)$_{ij}$*(Female)$_{ij}$ + b_{11} (EarlyHealthProblem)$_{lj}$ *(Female)$_{ij}$ + u_{ij}. Robust standard errors, clustered by twin pairs in parentheses. ***significant at 1% **significant at 5% *significant at 10%

Appendix Table 5: Descriptive Statistics, BFRSS

	1987 BFRSS (1)	1994 BFRSS (2)	2000 BFRSS (3)
Background Characteristics			
Age	46.18 (14.83)	47.14 (5.64)	74.98 (3.45)
25-34	0.28 (0.45)		
35-44	0.24 (0.42)	0.38 (0.47)	
45-54	0.15 (0.35)	0.48 (0.50)	
55-64	0.16 (0.36)	0.14 (0. 34)	
65-75	0.17 (0.37)		0.58 (0.49)
76-85			0.42 (0.49)
Female	0.56 (0.50)	0.55 (0.50)	0.00 (0.00)
Grades of schooling	12.93 (3.14)	13.46 (2.65)	12.96 (3.04)
Less than 12 grades	0.17 (0.37)	0.09 (0.28)	0.17 (0.37)
12 grades	0.34 (0.47)	0.31 (0.46)	0.30 (0.46)
13-15 grades	0.24 (0.43)	0.28 (0.45)	0.22 (0.41)
16 grades or more	0.25 (0.43)	0.32 (0.47)	0.31 (0.46)
Self-reported Health		3.78 (1.05)	3.11 (1.12)
BMI	24.80 (4.26)	26.10 (4.83)	26.35 (3.83)
Overweight	0.43 (0.49)	0.55 (0.50)	0.63 (0.48)
Obese	0.10 (0.30)	0.17 (0.37)	0.15 (0.35)
Never Smoke	0.45 (0.50)		0.30 (0.46)

Notes: Standard deviations in parentheses.

Appendix Table 6: Differential Effect of Schooling in Twin and BFRSS Datasets

	Panel A: Combined MATR-BFRSS data set			
	BMI	**Overweight**	**Obese**	**Never Smoke**
schooling	-0.153	-0.015	-0.01	0.017
	(0.008)***	(0.001)***	(0.001)***	(0.008)***
twin data	1.04	0.125	0.048	0.019
	(0.433)**	(0.048)**	-0.033	-0.051
twin data*schooling	-0.088	-0.011	-0.004	0.002
	(0.031)**	(0.004)***	(0.002)*	-0.004
N	38153	38153	38153	37962

	Panel B: Combined MTR-BFRSS data set			
	Self-reported health	***BMI***	***Overweight***	***Obese***
schooling	0.1	-0.17	-0.013	-0.011
	(0.003)***	(0.012)***	(0.001)***	(0.001)***
twin data	1.31	0.301	0.008	-0.002
	-0.115	-0.724	-0.081	-0.061
twin data*schooling	-0.056	-0.029	-0.001	0
	(0.008)***	-0.05	-0.005	-0.004
N	29208	29220	29220	29220

	Panel C: Combined NAS-NRC-BFRSS data set				
	Self-reported health	***BMI***	***Overweight***	***Obese***	***Never Smoke***
schooling	0.077	-0.081	-0.007	-0.008	0.013
	(0.006)***	(0.019)***	(0.002)***	(0.002)***	(0.002)***
twin data	0.423	-0.176	0.03	-0.073	-0.184
	(0.133)***	-0.451	-0.062	-0.041	(0.060)**
twin data*schooling	-0.01	-0.028	-0.006	0.001	0.017
	-0.01	-0.031	-0.004	-0.003	(0.004)***
N	6232	6257	6257	6257	6254

Notes: Estimates are based on the regression: $Y_i = b_o + b_1 Schooling_i + b_2(twin\ data)_i + b_3(twin\ data*Schooling)_i + b_4 age_i + b_5 age^2_i + b_6 female_i + u_i$. Twin Data is a dummy variable equal to 1 if the observation is from the twin dataset. Robust standard errors in parentheses. ***significant at 1% **significant at 5% *significant at 10%