TI 2008-027/3
Tinbergen Institute Discussion Paper

The Health Returns to Education – What can we learn from Twins?

Petter Lundborg

Dept. of Economics, Free University Amsterdam, and Tinbergen Institute.

Tinbergen Institute
The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam
Roetersstraat 31
1018 WB Amsterdam
The Netherlands
Tel.: +31(0)20 551 3500
Fax: +31(0)20 551 3555

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Most TI discussion papers can be downloaded at http://www.tinbergen.nl.

The Health Returns to Education - What Can We Learn from Twins?

Petter Lundborg *
Department of Economics
Free University Amsterdam

March 10, 2008

Abstract

This paper estimates the health returns to education, using data on identical twins. I adopt a twin-differences strategy in order to obtain estimates that are not biased by unobserved family background and genetic traits that may affect both education and health. I further investigate to what extent within-twin-pair differences in schooling correlates with within-twin-pair differences in early life health and parent-child relations. The results suggest a causal effect of education on health. Higher educational levels are found to be positively related to self-reported health but negatively related to the number of chronic conditions. Lifestyle factors, such as smoking and overweight, are found to contribute little to the education/health gradient. I am also able to rule out occupational hazards and health insurance coverage as explanations for the gradient. In addition, I find no evidence of heterogenous effects of education by parental education. Finally, the results suggest that factors that may vary within twin pairs, such as birth weight, early life health, parental treatment and relation with parents, do not predict within-twin pair differences in schooling, lending additional credibility to my estimates and to the general vailidy of using a twin-differences design to study the returns to education.

JEL Classification: I12, I11, J14, J12, C41

Keywords: health production, education, schooling, twins, siblings, returns to education, ability bias

*Free University Amsterdam. Address: Free University Amsterdam, Department of Economics, De Boelelaan 1105, 1081 HV Amsterdam. Tel: +31 20 598 28 74. Email: plundborg@feweb.vu.nl. Personal homepage: http://staff.feweb.vu.nl/plundborg. Affiliated with Tinbergen Institute, IZA, HEP, and LUCHE.

1 Introduction

There is a long tradition in economics to estimate the private returns to schooling, as reflected through the effect of additional schooling on earnings. If schooling affects social well-being beyond its effect on earnings, however, these estimates will only partially capture the total returns to schooling. Evidence is now mounting that schooling is associated with several non-market outcomes, such as health, child's schooling and cognitive development, marital choices, fertility control, and crime (see Wolfe and Haveman 2002 and Grossman 2006 for recent overviews). Accounting for such outcomes may lead to different conclusions regarding the individual and social values of schooling.

Among these various non-market returns to schooling, there has recently been a growing interest in the health returns. Schooling is strongly associated with a range of different health measures and the relationship has been observed in many countries and time periods (Cutler and Lleras-Muney 2008). Since health care expenditures contribute only little to health in developed countries, the question arises whether education policies could be used to improve population health. The answer to this question hinges ultimately on whether schooling has a causal effect on health. Evidence on the issue is still scarce, however.

In this paper, my aim is to estimate the causal effect of education on health. In addition, I aim to explore some of the mechanisms through which the effect arises. I will base my estimates on a nationally representative sample of identical twins from the MIDUS survey in the US. Identical twins share common genes and, to a large extent, a common family background. By relating within-twin-pair differences in education to within-twin-pair differences in health, I am therefore able to difference out the influence of unobserved genetic traits and common family background that may otherwise bias the schooling coefficient. To the best of my knowledge, this is the first study using a twin-differencing approach to study the health returns to education.

In order to address the endogeneity of education, a number of recent studies have relied on various natural experiments, such as schooling reforms (see Grossman 2006 for an extensive overview). While these studies have certainly enhanced our understanding about the health returns to schooling, they rely on natural experiments that affect individuals whose return to schooling is likely to be different from the average returns in the population (Cutler and Lleras-Muney 2008). Changes in mandatory schooling laws, for instance, were typically intended to increase the schooling of those at the lower end of the education distribution, while having little or no effect on those planning to go to further studies anyway. Since the resulting estimates therefore reflect Local Average Treatment Effects (LATE), these studies tell us little about the effect on health of raising the schooling level for the entire population (ATE).[1] In addition, several of these recent studies, reviewed in greater detail in Section 2, face a problem of weak instruments, yielding imprecise and inconsistent estimates.

[1] Moreover, since the studies rely on different natural experiments, straightforward comparisons of the results are difficult, as different sub-populations are affected by the experiments.

Could a twin-differencing strategy be helpful in overcoming these shortcomings? Twin differences as a "natural experiment" relies on the existence of differences in schooling within identical twin pairs. If such differences only existed in particular types of twin pairs, a Local Average Treatment Effect would still be estimated. On the other hand, if the differences in schooling are equally distributed across twin pairs, the resulting estimate would come closer to reflect an Average Treatment Effect. Data from this paper and from Bonjour et al. (2004), for instance, seem to suggest the latter situation. Moreover, a twin-differences strategy does not rely on natural experiments that are often only weakly related to schooling levels. A twin-differences strategy therefore has the potential to provide new and important knowledge about the health returns to education in the population.

While a twin design have some distinct advantages, it also brings problems of its own. The main criticism of twin studies has been that while twin differencing will remove the influence of unobserved factors common to a twin pair, there may still remain within-twin-pair differences in unobserved factors that affect schooling. Bound and Solon (1999) showed that any ability differences within twin pairs that are not removed in a twin-fixed-effects model could potentially increase the endogeneity bias compared to OLS estimates.[2] As a major candidate for such within-twin-pair differences, Bound and Solon (1999) mention birthweight, since some evidence suggests that low birthweight may be correlated with ability and early life health. While even identical twins may differ in birth weight, there is to date mixed evidence as to whether such differences are also associated with within-twin-pair differences in schooling (see e.g. Behrman and Rosenzweig 2004; Miller 2005).

Besides birth weight and early life health, there may exist within-twin-pair differences in other unobserved factors as well, such as parental treatment and relation to parents (Ashenfelter and Rouse 1998). Ashenfelter and Rouse (1998) provided suggestive evidence that parents try to treat their twins as equally as possible, finding, for instance, that parents commonly give their twins similar names, names that rhyme, or names starting with the same letter. The evidence is still limited, however, and to the extent that differences in parental treatment are also related to differences in schooling, twin-fixed-effects estimates of the health returns to schooling may still be upward biased.

In this paper, I am able to address these issues using unique and detailed information on within-twin-pair differences in factors such as early life health, birthweight, classroom placement, peer choices, and parent-child relations. This allows me to investigate the importance of several of the commonly cited factors that might give rise to endogenous schooling differences within twin pairs. Following the approach of Ashenfelter and Rouse (1998) and Bonjour et al. (2004), I will first estimate the correlation between average twin-pair education and average twin-pair early life characteristics that may be correlated with "ability", such as birthweight, early life mental and physical health, and parent-child

[2] To see this, first note that the ability bias is determined by the ratio of exogenous variation to total variation. *If* differencing reduces the fraction of exogenous variation, ability bias may increase.

relations. This will give me an indication of the expected "ability" bias in the regressions. I will then compare these estimates to those obtained from regressions on within-twin-pair differences in education on within-twin-pair differences in the same early life characteristics. The latter will indicate the ability bias in the within-twin-pair regressions. A comparison reveals the extent to which the "ability" bias is removed in the within-twin-pair regressions.

Previous studies, such as Ashenfelter and Kreuger (1994) and Ashenfelter and Rouse (1998), did not have access to such detailed information on early life conditions of the twins. With this data, I am therefore able to check the credibility of my estimates and also provide more general insights into the credibility of using a twin-differences strategy.

Finally, twin-differencing raise the issue of measurement errors in reported schooling. If individuals misreport educational attainment, such errors are exacerbated by differencing, and even more so when differencing between identical twins, causing the estimate of schooling to be downward biased (Griliches 1979). The typical solution, proposed by Ashenfelter and Kreuger (1994), has been to instrument for schooling, using a co-twins report on one's own schooling. This issue is important in the literature on the wage returns to schooling, where assessing the exact magnitude of the returns is central. In the health returns literature, it still remains to settle whether or not there exists a causal effect *at all*. While I am able to address the issue of within-twin pair differences in factors potentially associated with within-twin pair differences in schooling, the data does not allow me to adress the measurement error problem by instrumenting. For my purposes, however, is it more important to address the potential upward bias in the results than any downward bias caused by measurement errors. I will, however, make use of previous estimates in order to get an idea of the likely downward bias in the estimates.

My results suggest a causal effect of education on health. Higher educational levels are found to be positively related to self-reported health but negatively related to the number of chronic conditions. In contrast, I find no evidence that important lifestyle factors, such as smoking and obesity, contribute to the education/health gradient. I am also able to rule out occupational hazards and health insurance coverage as explanations for the gradient. Finally, my results suggest that factors that may vary within twin pairs, such as birth weight, early life health, and parent-child relations, do not predict within-twin pair differences in schooling, lending additional credibility to my estimates and to the twin-differencing design in general.

I start the paper by giving some background to the education/health literature and discuss some recent findings. I then discuss the data and compare it to data from CPS in order to assess its generalisability. Next, I discuss the empirical model. I then report the results, where the results from the pooled twin sample are contrasted with those obtained when applying a twin-differences strategy. Finally, the results are discussed and some conclusions are drawn.

2 Background

There are basically three ways in which the link between education and health has been explained. First, education may make people more efficient in producing their own health, suggesting a causal effect running from education to health. This is how education enters the demand-for-health model (Grossman 1972). Here, educated people obtain a larger health output from a given amount of health inputs. Schooling may also increase allocative efficiency in the production of health (Rosenzweig and Schultz 1982; Kenkel 1991). In this case, educated people are able to pick a better mix of inputs in the production of their own health.

Second, education and health may be related through unobservables, such as family background and genetic traits. Fuchs (1982) proposed time preferences as such an unobserved variable, where less future-oriented people will invest less in both education and health than more future-oriented people, since the benefits of such investments are of long-run character. Labelling all such kind of unobserved factors "ability", its omission in a regression will bias the coefficient of schooling.

Third, health may affect educational attainment. Poor health early in life may intervene with learning and schooling choices and may also be associated with health later in life. Some evidence suggests, for instance, that low birthweight, being an early health marker, is associated with less schooling being obtained (Behrman and Rosenzweig 2004; Black et al. 2007).

The discussion above suggests that schooling may be endogenous, leading to inconsistent coefficient estimates of schooling. To deal with this, early studies used various instruments, such as parental education as instrument for own education (Berger and Leigh, 1989; Sander, 1995; Leigh and Dhir, 1997). The earliest study, Berger and Leigh (1989) used IQ, per capita income and per capita expenditures on education in the state of birth, and parents' schooling as instruments. For several reasons, the exogeneity of these instruments may be questioned. IQ may be correlated directly with health and parents' schooling may be correlated with child health, which in turn affects later life health. Moreover, state income and education expenditures may be correlated with health expenditures and other state characteristics that affect health.

The exclusion restrictions are easier to defend in a number of recent studies, utilising various "natural experiments" to estimate the effect of schooling of health. Grossman (2006) identifies six such studies (Adams 2002; Spasojevic 2003; Arkes 2004; Arendt, 2005; Lleras-Muney, 2005; de Walque 2007). Since then, an additional number of studies using natural experiments have appeared, such as Kenkel et al. (2006), Oreopoulos (2006), Grimard and Parent (2007), and Chou et al. (2007).[3] In comparison with the extensive literature on the wage returns to education, the evidence base is still small, however.

Five of the studies use educational reforms as a mean to identify the effect (Spasojevic, 2003; Lleras-Muney, 2005; Arendt, 2005; Oreopoulos 2006; Chou

[3] Chou et al. (2007) consider the effect of parental education on child health.

et al. 2007). The success of this strategy has varied, however, since, among other things, several of the studies face a problem of weak instruments, leading to inconsistency of the IV-estimator. Arendt (2005) uses schooling reforms in Denmark in 1958 and 1975 that affected the entire population. This makes it difficult to distinguish between cohort effects and the influence of the schooling reform. Moreover, the study suffered from weak instruments and when instrumenting for education, Arendt (2005) does not find any significant effect of education on health. This is similar to Lleras-Muney (2005), who also faced a problem of weak instrument when using individual-level data.[4] Spasojevic (2003) only found a significant effect of education on health when applying one-tailed tests.

Similar problems of weak instruments exist in studies using other types of natural experiments to study the health/education gradient. Adams (2002) adopts the strategy of Angrist and Kreuger (1991), using quarter of birth as an instrument for education. The F-values on the instruments are only about 1, indicating a problem of weak instruments. Unsurprisingly, no significant effect of education on health was obtained.

As already mentioned in the introduction, these studies identifies local average treatment effects (LATE).[5] de Walque (2007), for instance, used the fact that college enrollment was one way to avoid being drafted for the Vietnam war. Risk of induction is then used as an instrument for going to college.[6] This means that the effect of education is only estimated for the subgroup of males that decided to go to college in order to avoid getting drafted. These individuals return to schooling may very well be different from the average returns in the population and the estimates thus represent local average treatment (LATE) effects.

The studies discussed above commonly find that instrumenting for education increases the magnitude of the education effect, although the estimated effects in many cases are not significant. The increase in magnitude has been explained in two ways. First, the instruments are based on policy interventions that affect the educational attainment of people only at the lower end of the education distribution. The returns to education for this group is likely to be greater than for the population in general. Second, random measurement errors in the schooling variable lead to a downward bias in the OLS estimates. Instrumenting for schooling may help remedy this problem, as long as the instruments are not correlated with the error (Card 2001).

To summarize; while recent studies have provided new and interesting findings on the effect of education on health, there still exists a great deal of un-

[4] She found a significant and positive effect of education on health when using data on synthetic cohorts and instrumenting with state-level schooling reforms, though. Mazumder (2007) finds, however, that the results are not robust to the inclusion of state-specific state trends.

[5] The exception is Oreoupoulos (2006), who claims that his estimate comes close to reflecting an Average Treatment Effect. He exploits the change in minimum school leaving age from 14 to 15 in the United Kingdom that affected half of the population of 14-year-olds

[6] de Walque (2007) also uses an alternative instrument, indicating the risk of induction times the risk of being killed in the war.

certainty about the causal effect. In addition, many of the studies suffer from low precision and identification being based on specific subgroups of the population, calling into question the representativeness of the results. While a twin-differences design is no panacea, it does avoid some of these problems and may therefore potentially bring interesting new findings to the literature.

3 Data

I base my estimates on data from the first wave of the Midlife in the United States (MIDUS) survey. The first wave collected data in 1995 on a total of 7,108 individuals. This baseline sample was comprised of individuals from four subsamples: (1) a national RDD (random digit dialing) sample (n=3,487); (2) oversamples from five metropolitan areas in the U.S. (n=757); (3) siblings of individuals from the RDD sample (n=950); and (4) a national RDD sample of twins (n=1,914). To be eligible for the survey, participants had to be non-institutionalized, English-speaking, living in the United States, and aged 25 to 74. The study was orgininally set up with the purpose of investigating the role of behavioral, psychological, and social factors in understanding age-related differences in physical and mental health.

The response rate for the telephone interviews in the first wave of MIDUS was 70%. Among these, 86.3% also completed a self-administered questionnaires (SAQ), giving an overall response rate of 60.8%.

The twin sample consists of 1,914 twins, participating in the MIDUS Twin Screening Project. The recruitment of the twins followed a two-stage sampling design. In the first stage, a representative national sample of approximately 50,000 households was screened to identify families with twins. Respondents were asked whether they or any of their immediate family members were members of intact twin pairs. In 14.8% cases, the respondent reported the presence of a twin in the family. These respondents were then asked whether the research team was allowed to contact the twins in order to solicit their participation in the survey. About 60% of the respondents agreed and where subsequently enrolled in the MIDUS recruitment process.

Second, twin households were contacted and offered to participate in the MIDUS survey. Twins that agreed to participate were asked to provide contact information for the co-twin. In a smaller number of cases, several twin pairs per family existed. To be included in the MIDUS twin study, the respondent through which twins were identified had to be related to the twin by being a spouse or partner, a sibling, a child (also for the spouse or partner), or a father or mother. Moreover, the twins had to be between the ages 25 and 74. Both twins also had to have a residential phone number, excluding individuals living in prisons, nursing homes, and college dormitories etc. In addition, both twins had to live in continental US, both had to speak English, and both twins had to be mentally and physically able to do the interview.

Applying these eligibilility criteria, almost half (49%) of the identified twin pairs were ineligible for the survey. The major reason (52%) was ineligibilility due to the age criteria. The second single most important reason was that the main respondent was not related to the identified twin according to the eligibility criteria (30%). A further 25% did not lead to completed interviews for various reasons. The most common reason (41%) was that the interviewer was unable to reach the twin or contact person, whereas the second most important reason (32%) was that the twin or contact person refused to participate.

It should be noted that MIDUS was the first national sample of twins that was ascertained randomly via telephone. Using nationally representative data is an improvement compared to prior economic studies using twin data, such as Ashenfelter and Kreuger (1994) Ashenfelter & Rouse (1998). These studies used highly selective data, collected during the Twinsburg twins festival. As noted by Ashenfelter and Rouse (1998), the twin pairs attending this festival may be more alike than a random selection of twin pairs, since the festival emphasises the similarity of the twins and the pairs attend in similar clothes and hairstyles.

By using information collected as part of the initial twin screening questionnaire, twin pairs were diagnosed as identical or fraternal twins. The questions used in the diagnosis included, for instance, whether the twins had the same eye color, natural hair color, and complexion, whether individuals mistook them for each other when they were young, and whether they had ever undergone testing or been told by a doctor whether they were genetically identical or fraternal. Based on their answers to the questions, the twins were assigned points, which were subsequently totaled. "High" scores indicated identical twin pairs and "low" scores indicated fraternal twin pairs. In a small number of cases, the pair's score fell in the middle of the range and no diagnosis was given. This method of diagnosing twin zygosity has proven reliable and has shown to be over 90% accurate in diagnosing twin zygosity (e.g., Nichols and Bilbro, 1966).

Out of the 1,996 twins, 32 twins were dropped due to uncertainty regarding zygosity. Of the remaining twins, 734, or 37%, were identical twins, which were then selected for the analysis. I dropped 3 twins who had yet not finished their education. In addition, I dropped 19 twins were id-number was lacking and 18 twins were information on the co-twin was lacking. This resulted in a final sample size of 694 identical twins.

3.1 Explanatory variables

Educational attainment was measured in 12 categories in MIDUS, ranging from no school/some grade school to PhD. For my main analysis, I categorized this variable into four categories, ranging from highest to lowest: at least a college degree; some college but less than a BA degree; a high school diploma; less than a high school diploma.

While years of schooling has been a common measure in many prior studies on the wage returns of education, I have several reasons for using educational categories instead as my main measure. First, it is not straightforward how to impute the years of schooling from these categories. Since measurement errors would inevitably increase from such as a procedure, this would accentuate the measurement error problem, which is already serious when taking twin differences. Second, the educational degree may be as relevant, or even more relevant, as years of schooling. In de Walque (2007), for instance, there is a sharp increase in the effect of number of years of schooling on smoking, once reaching college. Similar evidence for non-linear effects have been obtained in the literature on the wage returns to education (Hungerford and Solon 1987; Belman and Haywood 1991; Isacsson 2004). Based on such findings, some economists argue that credentials matter more than years of schooling (for a discussion on this, see Card 1999). I will also, however, provide some estimates based on imputations of years of schooling, using the results from Jaeger (1997).

In the regressions, I use the category less than a high school diploma as the omitted reference category. Besides education, the regressions control for age, gender, marital status, race, and total household income. The latter was obtained by summarising all sources of income for all members in the household.

3.2 Health outcomes

My two main measures of health is self-assessed health and the number of chronic conditions. The former was assesed through the following question: "Using a scale from 0 to 10 where 0 means "the worst possible health" and 10 means "the best possible health," how would you rate your health these days? Self-assessed health has been found to be a strong predictor of subsequent mortality (see for instance Idler and Benyiamini 1997). There are some concerns, however, about the interpretation of the responses. Older individuals often report similar self-reported health as younger persons, despite have "objectively" worse health (Groot 2000). I will therefore also consider a more "objective" health measure, measuring the number of chronic conditions.

Besides health measures, I will also examine lifestyle, occupational hazards, and health insurance coverage as outcomes variables. For the former, I use information on smoking, Body Mass Index, and physical exercise. Physical exercise is measured through the number of occasions during past month that the individual engages in vigorous physical activity. Occupational hazards were measured through two questions asking the respondent to rate the extent to which his/her job affected his/her health and about the number of work accidents during the past 5 years. Health insurance coverage was assessed by asking the respondent if he/she was covered, either through him/herself or through a spouse.

3.3 Representativeness of the sample

Next, I consider to what extent the sample of identical twins resembles the main MIDUS sample and the US population in general. Even though the aim of MIDUS was to obtain nationally representative samples for both the main sample and the twin sample, dropouts may affect the representativeness. For the latter purpose, I will make some comparisons with data from the Current Population Survey (CPS) of 1995. In Table 1, descriptive statistics for the three samples are shown.

A comparison between the MIDUS main sample and the twin sample reveals that the twins are significantly younger, are more likely to be white, are more likely to be married or cohabitating, have a higher income, have better health, are more physically active, and have a lower Body Mass Index compared to the main sample. There are no significant differences in the level of education, however. Moreover, there are no significant differences in the smoking rate, the overweight rate, the fraction holding a health insurance, or the fraction having experienced job hazards.

The comparison with the CPS data reveals that the both the twin sample and the MIDUS main sample are better educated than the US population in general. Similar patterns were found in several previous studies, reflecting a selection of better educated twins into the surveys (Ashenfelter and Kreuger 1994; Ashenfelter and Rouse 1998; Bonjour et al. 2003). While similar in terms of gender distribution, the twin sample also contains more whites and has a slightly more compressed age distribution than the CPS sample. Regarding marital status, the CPS from 1995 does not contain a straightforward estimate of the number of cohabitating or married couples. Considering marriage alone, however, the fraction of married in CPS in 1995 was 67.5%, compared to 71.6% in the twin sample and 62.6% in the MIDUS main sample.

The higher education of the twin sample has one implication; if the health returns to education exhibit diminishing returns to scale, I will most likely provide conservative estimates.

4 Empirical strategy

In this section, I describe the empirical strategy, based on twin-differencing. To see how this strategy may help us estimate the causal effect of education on health, first, consider an individual i, whose health stock H_i is determined by:

$$H_i = \beta S_i + \alpha A_i + u_i, \tag{1}$$

where S_i denotes schooling, A_i denotes unobserved "ability", and u_i is an unobserved random component. In this context, ability is taken to mean both unobserved genetic traits affecting health, as well as unobserved family background. Next, let schooling be determined by

$$S_i = \delta A_i + \xi_i, \tag{2}$$

where A_i denotes the same unobserved "ability" components that affects health and ξ_i denotes a schooling-specific random term.

This gives the standard result that an OLS estimate of β is biased such that:

$$p\lim(\beta_{OLS}) = \beta + \alpha\frac{\sigma_{AS}}{\sigma_S^2}. \tag{3}$$

Since unobserved ability is likely to be positively correlated with both schooling and health, is it usually assumed that an estimate of β_{OLS} will be upward biased.

Next, I turn to the twin-differencing strategy. Now, let H_{1j} and H_{2j} denote the health of the first and second twin in the jth twin pair. The unobserved component is again made up of two parts. The first part, μ_j, denotes unobserved factors that vary between twin pairs but not within pairs. This could, for instance, be genetic characteristics and early life environmental factors. Finally, ε_{1i} and ε_{2i} denote unobserved factors specific to each twin. This can be written as:

$$H_{1j} = \beta S_{1j} + \mu_j + \varepsilon_{1j}, \tag{4}$$

$$H_{2j} = \beta S_{2j} + \mu_j + \varepsilon_{2j}, \tag{5}$$

Next, I take the difference between (4) and (5), giving:

$$H_{1j} - H_{2j} = \beta_{WTP}(S_{1j} - S_{2j}) + \varepsilon_{1j} - \varepsilon_{2j}, \tag{6}$$

where β_{WTP} is the within-twin-pair estimate of education. In this specification, all factors that are common to both twins in a given twin pair will be differenced out. Since twins share common genes, their influence will vanish, as well as the influence of common family background. This means that an OLS estimate of (6) will no longer be biased due to unobserved twin-pair specific variables. Any remaining unobservables that remain in the error term after differencing may still, however, bias the results, if these unobservables are still related to both schooling and health. I will return to this issue in more detail in Section 4.4.

It is well known that measurement errors in schooling are exacerbated by differencing and even more so when differencing between identical twins. (Griliches 1979). This will cause twin FE estimates to be downward biased. The extent of downward bias due to measurement error may be calculated in the case where one has a measure of the reliability of self-reported schooling and a measure of the correlation in schooling within twin pairs. As shown by Griliches (1979), in the presence of classical measurement error, the twin FE estimate is then biased according to:

$$\beta_{WTP} = \left(1 - \frac{\mathrm{Var}(\nu)}{[\mathrm{Var}(S)]\,(1\text{-}\rho_{\mathrm{S}})}\right),$$

where $\mathrm{Var}(\nu)$ denotes the assumed common variance of the twins measurement error, $\mathrm{Var}(S)$ is the variance in the true schooling levels, and ρ_{S} is the correlation between the measured schooling levels of the twins. The part $\frac{\mathrm{Var}(\nu)}{\mathrm{Var}(S)}$ is called the reliability ratio. Research suggests that the reliability of self-reported schooling is typically about 90 percent, a figure that has been remarkable stable across studies (Card 1999). Moreover, the correlation in schooling within identical twin pairs is commonly found to be about 0.75 (see e.g. Ashenfelter and Rouse 1998). Taking these estimates together, an attenuation bias of about 30% is typicaly obtained.[7]

To obtain an estimate of the reliability ratio, previous studies have exploited data where several measures of the education of the respondent are given. Often, this has been a measure given by a co-twin (see e.g. Ashenfelther and Rouse 1998). Isacsson (1999), however, used a second measure on the respondent's education, taken from register data. While I do not have data on the co-twins report on the other twins education or access to register data, I do have a second measure of the respondent's education at the follow-up survey in 2004. The correlation between these measures suggest a reliability ratio of 0.90, being very much in line with previous estimates.[8] This is under the assumption, however, that the measurement errors of the two measures are uncorrelated, which is a strong assumption. If the measurement errors are positively correlated, the reliability ratio is overestimated. On the other hand, it should be noted that for some individuals, there may have been real changes in educational attainment between the waves, suggesting some downward bias in the reliability ratio. The estimated correlation in schooling within twin pairs in MIDUS is 0.72, which is also rather similar to the figures obtained in previous studies. Taken together with the estimated reliability ratio, this indicates that the twin FE estimator is biased downward by about 36%. Assuming reliability ratios of 0.85 or 0.95 instead, the downward bias would be 53% and 18%, respectively.

5 Results

5.1 Self-reported health

In Table 2, I show both OLS and twin FE results for self-reported health. Starting with the MIDUS main sample, the results show a strong and positive assocation between education and health. These results are largely mirrored in the pooled twin sample, with the difference that the magnitude of the associations between education and health are now somewhat increased. Having at least a college degree, for instance, is now associated with almost a one unit increase on the health rating scale compared to having no or at most some high school education. Having a high school diploma, compared to lowest category, is associated with a 0.5 unit increase on the rating scale.

[7] The classic solution the measurement error problem in twin studies, proposed by Ashenfelter and Kreuger (1994), has been to instrument for ones own education, using the co-twins report on ones own education. Since the MIDUS does not contain such information, I am not able to apply this strategy.

[8] Not all twins participated in the follow-up survey and my estimate is based on a sample of 541 identical twins.

In the third column, the results from the twin FE estimation are shown. Interestingly, the magnitude of the associations between the educational categories and self-reported health increase in the twin FE specification. For the two first educational categories, the magnitudes are almost doubled, whereas for the highest category the magnitude increased by about 35%. These results are somewhat surprising, since one would expect a weaker relationship, once the influence of genes and family background common to the twins are controlled for. Moreover, measurement errors are excaggerated using twin FE, suggesting that these twin FE estimates are downward biased by about 35%. It should be noted, however, that the confidence intervals of the estimates overlap to a great extent. In sum, the results suggest a strong effect of education on self-reported health that seems to increase in magnitude when controlling for genetic factors and common family background of the twins.

For imputed years of schooling, the OLS estimate suggest a small but significant and positive effect of schooling on self-reported health in the pooled twin sample (not shown here, but available on request). One additional year of schooling is associated with a 0.079 increase in self-reported health. The twin FE point estimate is rather similar; 0.067. This somewhat lower point estimate might also reflect downward bias due to measurement error in schooling that is exacerbated by differencing, however. The twin FE estimate of schooling is rather imprecisely measured and the coefficient is not significant.

Besides education, it is interesting to note that income shows a positive and significant effect on self-rated health in all three specifications. Controlling for twin-pair specific unobserved heterogeneity does not seem to reduce the magnitude of the income effect, which is remarkable stable across the specifications.

5.2 Chronic conditions

Next, I consider the association between education and the somewhat more objective measure of health; the self-reported number of chronic conditions. Columns 4-6 of Table 2 show the results for the three samples.

In the MIDUS main sample, education shows a strong and negative association with the number of chronic conditions, the association being strongest for the highest education category. In the latter case, having at least a college degree is associated with a decrease in the number of chronic conditions by 1.2. Income is again significant and is associated with a decrease in the number of chronic conditions.

In the pooled twin sample, education again shows a significant and negative association with the number of chronic conditions. The magnitude of the associations are greater than the corresponding ones in the main sample, with the two highest education categories now being associated with a decrease in the number of conditions by 2. Income is no longer significant and being white is now associated with a decline in the number of conditions.

The twin FE estimates tell a similar story. The significant associations between education and number of conditions remain for all education categories, except for the highest one. For the highest education category, the point estimate is still negative, though. The point estimates are lower than the corresponding ones obtained in the pooled twin sample and more close to the ones obtained for the MIDUS main sample. It should be remembered though that the twin FE estimates are most likely downward biased due to the measurement error problem. Taking this fact into account brings the estimates more in line with the ones obtained in the pooled twin sample.

Imputed years of schooling show a significant and negative association with chronic conditions in the pooled twin sample. One additional year of schooling is associated with a 0.09 decrease in the number of chronic conditions. In contrast, the twin FE estimate of schooling suggest a small and positive relationship between schooling and the number of chronic conditions. This estimate is imprecisely measured, however, and not significant.

5.3 Investingating the mechanisms

In order to examine potential mechanisms through which education affects health, I will next investigate the effect of education on various lifestyle factors, occupational hazards, and health insurance coverage. Since smoking and overweight are the two main causes of preventable deaths in the US, I will focus on these lifestyle factors. In addition, I will consider physical activity, since it relates to overweight and has other health benefits as well. In order to preserve space, I will from now on only compare the results obtained from the pooled twin sample with the twin FE estimates.

5.3.1 Smoking

While a number of studies have found a negative correlation between smoking and education, there are reasons for interpreting these results with some caution (see, for instance, de Walque 2007 and Grimard and Parent 2007). First of all, smoking is usually initiated before schooling is completed, suggesting that part of the effect of education on smoking may be explained by unobserved third factors or reverse causality running from smoking to education. Second, the dangers of smoking are well known and several studies show that people in general overestimate the risks (Viscusi 1990; 1991, Lundborg and Lindgren 2004, Lundborg 2007). If anything, more educated people should hold risk perceptions more closely related to the actual risks, suggesting that education should be associated with lower risk perceptions. So, if the assocation is mainly due to unobserved factors affecting both schooling and smoking, such as time preferences, and if these factors are common to twins, we would expect the effect to vanish when employing twin FE.

Starting with the pooled twin sample, the first column of Table 3 shows a strong association between education and smoking, that increases with the level of education. In contrast, the twin FE estimates of education are just between one fifth and half the magnitude and insignificant in all cases. While this is consistent with there being a substantial downward bias in the results, it is also consistent with the hypothesis that unobserved factors, such as genetic traits or time preferences are driving the results for the pooled twin sample.

To further investigate the issue, I re-ran the regressions, this time replacing the smoking measure with a measure of smoking at the age of 16 or earlier. For obvious reasons, education should not have any causal effect on smoking at age 16 or earlier. A shown in Table 3, however, there is a strong and significant negative correlation between having at least a college degree and smoking at 16 or earlier in the pooled twin sample. This is in sharp contrast with the twin FE results, where the effect of having at least a college degree is just one tenth of the effect obtained in the pooled twin sample and not signficant. In sum, these results provide no evidence that the causal effect of education on health runs through smoking behaviour.

Does the results change when using imputed years of schooling as a measure of education? A significant and negative effect of schooling on smoking is obtained in the pooled twin sample. Here, one additional year of schooling is associated with a 0.04 percentage points decrease in smoking. The twin FE point estimate is half in magnitude and not significant. In sum, these results are in line with the results using educational categories as measure of education.

5.3.2 Physical activity and overweight

Next, I investigate the association between education and physical activity and Body Mass Index. Recent evidence from Kenkel et al. (2006) suggests a causal link between education and physical activity and overweight. Since MIDUS contains several measures of physical activity, I opted for the ones that are most likely to reflect deliberate attempts to be physically active, such as being physically active during the winter.

In the first two columns of Table 4, I show the association for the pooled twin sample and the results from the twin FE estimation. In the pooled twin sample, having some college or having a college degree is associated with about two more occasions of physical activity per month compared to the reference category. Having graduated high school shows no signficant effect. Surprisingly, the results get even stronger when employing the twin FE estimator. Now all the educational categories are associated with an increase in the number of occasions of physical activity by about 3. The educational categories are all significant at least on the 10% level.[9]

To investigate to what extent the higher physical activity of educated individuals also transforms itself into lower body mass and a lower prevalence of overweight and obesity, I next examine the direct association between education and these outcomes. Column 3 to 6 of Table 4 shows the results for the pooled twin sample and the results from the twin FE estimator. As shown in

[9] Similar results were obtained when using alternative measures, such as moderate activity during the summer and vigorous activity during the winter. Only for vigorous activity during the summer were the results from the twin FE not significant. The point estimates were, however, rather similar to those from the pooled twin sample, where the two highest educational categories were positive and significant at the 10% level.

column 4 and 6, in the pooled twin sample, education shows a strong and negative significant correlation with both BMI and overweight for all educational categories. Belonging to the highest educational category is for instance associated with a 3.2 decrease in BMI compared to the omitted reference category. These effects are completely swept away in the twin FE estimates, however. The point estimates of education are now in most cases only a tiny fraction of those obtained from the pooled twin sample and are no longer significant. For instance, belonging to the highest educational category is now associated with a 0.02 increase in BMI, with a p-value of 0.98. Assuming a downward bias in the twin FE estimates by about 36%, the difference in point estimates seems too large to be explained by measurement error in schooling alone.[10] The results using imputed years of schooling largely mirror these results. While years of schooling is significantly and negatively related to BMI in the pooled twin sample, the twin FE point estimate is a lot smaller and not significant.[11] For overweight, years of schooling is insignificant in both specifications. In sum, the results suggest that while there seems to be causal link between education and physical activity, this does not transform itself into a causal effect of education on body mass.

5.3.3 Occupational hazards

Another potentially important explanation for the education/health gradient is that educated people are able to obtain less risky jobs. While previous studies have suggested that job risks explain little of the education/health gradient, these studies have not been able to control for some of the endogeneity of job risks (Lahema et al. 2004).

As shown in Table 5, education shows no significant association with neither measure of job risks in the pooled twin sample. There results are mirrored in the twin FE estimates.[12] An an additional check, I also re-ran the regressions on self-rated health and chronic conditions, this time controlling for job risks.[13] While job risks showed a significant and negative correlation on self-rated health in the pooled twin sample, the coefficients of education are almost unchanged in comparison with the estimates shown in Table 2 for the pooled twin sample. In contrast, the twin FE estimates showed no significant association with job risks, while the effect of education was still positive and significant. For chronic conditions, job risks did not show any statistically significant association in either specification. In sum, no evidence is obtained that job risks are an important explanation for the health/education gradient.

5.3.4 Health insurance coverage

[10] Similar results were obtained when using alternative measures, such as waist-to-hip ratio and obesity.

[11] Not shown here, but available on request.

[12] The results were similar using imputed years of schooling.

[13] Not shown here, but available on request.

Can differences in health insurance coverage explain the health/education gradient? In general, health insurance coverage has been found to have only small effects on health (Cutler and Lleras-Muney 2008). Moreover, in the MIDUS survey, about 90% of the respondents are covered by health insurance, making it unlikely to be a main driver of the education/health gradient. In order to examine this, however, Table 6 shows the correlation between education and health insurance coverage.

In the pooled twins sample, education clearly seems to be associated with a greater likelihood of being covered by health insurance. The twin FE estimates tell another story, however, suggesting that the relation is non-causal.[14] Interestingly, the point estimates of the education categories are now negative, although being insignificant. To further investigate matters, I re-ran the regressions on self-reported health and the number of chronic conditions, this time also including health insurance as a covariate.[15] The coefficients of the education variables were virtually unchanged in all specifications. In the pooled twin regression, health insurance showed a *negative* association with self-reported health that was significant at the 10% level. The twin FE estimates showed no significant association between health insurance and health, however. In sum, the results suggest that the correlation between education and health insurance is most likely driven by unobserved factors.

5.4 Heterogenous effects

One of the explanations for the general finding that IV estimates often exceed their OLS counterparts, when using schooling reforms to estimate the effects of education on wages or health, is that reforms usually only lower socio-economic groups, where the returns to education are greater than for the general population. In this section, I will investigate whether the health returns to education differs by parental education.

I interacted each differenced education dummy with a 1-4 measure of average parental educational attainment.[16] The results for self-reported health, shown in Table 7, suggest that the health returns to education decline with the level of parental education. It should be noted, however, that only the interaction between having a high school degree and parental education is significant. In the case of chronic conditions, none of the interaction effects are significant in either specification. In sum, these results provide only little support that the health returns to education differ by parents' education.

I also tested whether the estimated effects of education varied by age and gender. In the pooled twin sample, there was a trend towards a decreasing effect

[14] Similar results were obtained when using imputed years of schooling.

[15] These results are not shown in the Table, but are available upon request.

[16] The 1-4 variable measures the same educational categories as for the main respondent. In order to address measurement error in parents' education, I followed the approach of Ashenfelter & Kreuger (1994) and averaged the twins' reports before creating the variable. In cases where only one of the twins reported, I used that measure.

of education on health for higher ages for both self-reported health and chronic conditions. The interaction terms were not significant, however, and the pattern was not obtained in the twin FE estimates. No differential effect of education by gender was obtained in any of the specifications.

5.5 Differences within twin pairs

My twin FE estimates may still be biased, if there are individual-twin-specific factors that are not removed by differencing and that determine within-twin-pair differences in schooling. In this section, I will investigate this issue, by relating witin-twin-pair differences in various potentially important early life conditions to within-twin-pair differences in education. First, however, I will give some descriptives about early life differences between twins, summarised in Table 8.

First of all, it is of interest to examine the extent to which parents treat twins similarly. If parents, for instance, treat a less able twin differently from a more able twin, this may affect schooling choices and later life health, potentially biasing the twin FE estimates. Suggestive evidence is given in Ashenfelter and Rouse (1998), where twins are found to be given similar names with a frequency that is much higher than what would be expected by chance alone. However, anecdotal evidence also suggest that parents may try to emphasize the differences between the twins, for instance by dressing them differently or giving them different haircuts. Such differential treatment may be a potential source of educational differences and possibly also give rise to health differences. Data from MIDUS does not, however, support the latter kind of parental behaviour. In the first row of Table 8, the results from a question about how often their parents, or the people who raised them, did things like dress them differently or give them different haircuts are shown. The answer was Never in 85% of the cases. In only 8% of the cases was the answer most or all of the time. This provides some suggestive evidence that parents try to treat twins similarily.

Another choice that parents face is whether or not to put their twins in the same school and/or the same class. As shown in the second row of Table 8, however, the majority of parents prefer to keep twins in the same class, as 57% answered that they were in the same classroom always or most of the time. In only 14% of the cases was the answer Never. A relevant question is also whether twins who are separated at school more often end up with different educational attainment. If this is the case, and if classroom placement is largely random, this would be one source of exogenous schooling differences within twin pairs.[17] I obtain some evidence for this, since among the twin pairs being always or almost always in the same classroom, 39% end up with the same education, compared to 33% among those pairs reporting being sometimes or never in the same classroom. The difference is not statistically significant at conventional levels, however. Comparing only those being either always in the

[17] Another possibility would be that parents are more inclined to place twins with different abilities in separate classrooms. One reason for this would be to reduce the amount of competition between the twins. In this scenario, classroom separation would be endogenous.

same classroom with those being never in the same classrom reveals a somewhat larger differences, however; 39% vs 31%.

Another potentially important source of within-twin-pair differences in educational attainment and health is differences in their peer groups. Twins may self-select into different peer groups or face different peer groups due to classroom placement policies. Unobserved peer behavior that differ within twin pairs may therefore threaten the validity of twin FE estimates. This worry seems largely unfounded, however, since as many as 90% of the twins report that they always or most of the time had the same playmates. Only 2% report that they never had the same playmates.

Next, I turn to a more rigorous examination of between-twin-pair and within-twin-pair differences in various factors and their association with between-twin-pair and within-twin pair differences in educational attainment. I will start with one of the earliest within-twin-pair differences that can arise; differences in birthweight. Even though identical twins share common genes, the first born is usually heavier than the second born. This is confirmed in the MIDUS data, where the first born is on average 77 grams heavier than the second-born. Such differences may correlate with ability, cognitive functioning and later health and, thus, also with educational attainment. In MIDUS, I have complete information on the birth weight for 206 twins, or 104 twin pairs. Twins are generally lighter, something which is also confirmed in the data, with the average birth weight being 2,331 grams.

In the first column of Table 9, the correlation between average twin-pair education and average twin-pair birthweight is shown. The results in the table are based on a 1-4 measure of education, corresponding to the four categories used in the regressions above. I tried different measures of education, for instance binary indicators of high/low education, but the results did not change to any important extent. The between-twin-pair correlation in average birthweight and average education is positive but very small, 0.0001, and not signficant. I also tried an indicator of low birthweight, i.e. below 2,500 grams. This resulted in a small, insignificant, negative correlation, -.008. Finally, I also tried an indicator of having very low birthweight, i.e. below 1,500 grams, resulting in a negative correlation of -0.15 that was signficant at the 10% level. This does suggest that families with an average very low birth weight also have a lower average educational level, giving weak evidence for ability and family background affecting schooling.

The real question is, however, to what extent differences in birthweight within twin pairs affect within-twin-pair differences in schooling. The second column of Table 9, shows the correlation between differences in education within twin pairs and differences in birthweight within twin pairs. The correlation is again very small, 0.0001, and insignificant. Similar results are obtained for the indicators of low birth weight and very low birth weight. The latter shows a negative correlation with education, but is not significant (p=0.37). These results were similar estimating linear probability fixed effects models using binary indicators of schooling. If birthweight picks up some ability differences, these results suggest that between-twin-pair differences in education are more affected by ability differences than differences in education within pairs.

The remaining rows of Table 9, shows between-twin-pair and within-twin-pair correlations in other areas. An important source of differences in education obtained may be early life differences in health. To address this, I use measures of self-reported physical and mental health at age 16, which is given retrospectively by the respondents. These measures capture health differences that exist prior to completing schooling. The variables ranges from 1 to 5, where 1 denotes poor health and 5 excellent health. The second row of Table 9 shows the between-twin-pair and within-twin-pair correlation in early health and educational attainment.

The between-twin-pair correlation in average self-reported health at 16 and educational attainment is positive, but small and insignificant; 0.06. The within-twin-pair correlation is similar, 0.07, but again not significant. For mental health, the results are largely mirrored, the correlation being 0.04 both between twin-pairs and within twin pairs. This suggest that differences within twin pairs in self-reported physical and mental health at age 16, which could pick up ability differences, does not affect educational attainment and therefore do not bias my twin FE estimates.

Next, I will consider a range of indicators of parent-child relations. These measures reflect factors such as time and attention given by parents, love and affection given, strictness about rules, punishments, rating of relationship, parents expectations, and physical abuse.[18] The between-twin-pair correlation between average education and several of these factors is significant. For instance, the results show that the time and attention given by the mother is significantly and negatively related to educational attainment. Moreover, having a father who was less strict about rules shows a significant and positive association with education. Having a mother who held low expectations about the respondent is negatively related to educational attainment. Finally, having a mother or a father who beated or hit the respondent show a negative correlation with educational attainment.

None of these associations are significant in the within-twin-pair regressions, however. The only exception is the variable indicating how much love and affection the father showed. This variable is significant and actually shows a negative correlation with educational attainment in the within-twin-pair regression. To summarise, the within-twin pair differences in schooling are uncorrelated with almost all of these rather detailed measures of early life differences in early health and parental treatment. These findings support those of Ashenfelter and Rouse (1998) and Bonjour et al. (2004). These results lend some credibility to the results in this paper, as well as to the general validity of using a twin-differences design to study the returns to schooling.

[18] To assess the relationship to the father/mother, the respondent was asked to rate it on a 1-5 scale, were 1 means excellent and 5 means poor. For the other questions indicating parental treatment, the scale went to 1 to 4, where 1 indicates a lot and 4 not at all. Physical abuse was assessed by asking how often the mother/father pushed, grabbed, shoved, slapped, or threw something at the respondent. The scale went from 1 (often) to 4 (never).

6 Conclusion and discussion

I used a sample of identical twins to estimate the health returns to different levels of education. The results suggest a causal effect running from education to health. Higher educational levels are found to be positively related to self-reported health but negatively related to the number of chronic conditions. In contrast, estimates based on imputed number of years of schooling showed only small associations with health in the pooled twin sample and no significant association when employing twin FE methods.

My results do not provide any evidence that that the education/health gradient works through important lifestyle factors, such as smoking and overweight, or factors such as job risks and health insurance coverage. To the best of my knowledge, this is the first attempt to apply a twin-differencing approach to the topic. A twin-differencing approach may provide estimates that come closer to reflect an Average Treatment Effect compared to studies using educational reforms, for instance, to identify the effect of education on health. Further studies should continue to explore the mechanisms, while properly controlling for the endogeneity of education.

My results does not provide any evidence that unobserved "ability" differences within twin pairs are biasing my within-twin-pair estimates. I investigated this by first estimating the correlation between average twin-pair education and average average twin-pair early life characteristics that may be correlated with "ability" and/or time preferences, such as birthweight, early life mental and physical health, early health behaviours, and parental treatment. By comparing these estimates with those obtained from regressions on within-twin-pair differences in education on within-twin-pair differences in the same early life characteristics, I was able to get an indication of the expected "ability" bias in the regressions. The results indicated that the ability bias is less in the within-twin-pair estimates.

For self-reported health, I found that the twin FE estimates exceeded the OLS estimates. This is a bit unexpected, since it is usually assumed that the OLS estimates are upward biased and that controlling for unobserved ability will reduce the magnitude of the estimates. A similar results for the wage returns to education was obtained by Ashenfelter and Kreuger (1994).[19] One interpretation is that the correlation between ability, schooling, and health is more complex than what is usually assumed. If unobserved components, such as ability, affects the marginal cost of schooling, but not the marginal benefit, a negative correlation between ability and schooling may result. For instance, the marginal cost of schooling may be higher for people with high ability, since the foregone earnings are greater. If twin differencing removes unobserved ability, estimates will then increase in magnitude.

While I was able to address the issue of within-twin-pair ability bias, I was not able to account for the influence of measurement errors in the reports on

[19] This result did not hold, however, when Rouse (1998) and Ashenfelter and Rouse (1998) replicated the study with larger samples, suggesting that the finding of Ashenfelter and Kreuger (1994) was an artifact of their sample.

schooling. It should be noted, though, that ability bias give rise to an upward bias in the estimates, whereas measurement errors give rise to a downward bias. For the purpose of this paper, it was more important to address the former problem, since knowledge is still needed as to whether education has a causal effect *at all* on health.

In future work, I will consider a much larger sample of twins, drawn from twin registers. This will allow me to address the issue of heterogenous health returns to education with much greater precision, to adress the measurement errors problem, as well as to examine a greater range of health outcomes and health behaviours.

7 References

Adams, Scott. 2002. "Educational Attainment and Health: Evidence from a Sample of Older Adults." *Education Economics,* 10(20), 97–109.

Angrist, Joshua, and Alan Krueger. 1991. "Does Compulsory School Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics*, 106(4), 979–1014.

Arendt, Jacob. 2005. "Does Education Cause Better Health? A Panel Data Analysis Using School Reforms for Identification." *Economics of Education Review,* 24(2), 149–160.

Arkes, Jeremy. 2004. "Does Schooling Improve Adult Health?" Working Paper, RAND Corporation, Santa Monica, CA.

Ashenfelter, Orley, and Alan Krueger. 1994. "Estimating the Returns to Schooling Using a New Sample of Twins." *American Economic Review*, 84(5): 1157–1174.

Ashenfelter, Orley, and Cecilia Rouse. 1998. "Income, Schooling and Ability: Evidence from a New Sample of Identical Twins. *Quarterly Journal of Economics,* 113(1), 253–284.

Behrman, Jere R. and Mark R. Rosenzweig. 2004. "Returns to Birthweight." *Review of Economics and Statistics* 86, 586–601.

Belman, Dale and John Heywood.1991. "Sheepskin Effects in the Return to Education", *Review of Economics and Statistics* 73: 720-724.

Berger, Mark. C., and J. Paul Leigh. 1989. "Education, Self-Selection and Health." *Journal of Human Resources* 24(3): 433–455.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2007. “From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes.” *Quarterly Journal of Economics*, 122(1): 409-439.

Bonjour, Dorothe, Lynn Cherkas, Jonathon Haskel, Denise Hawkes and Tim Spector. 2003. “Returns to Education: Evidence from UK Twins." *American Economic Review*, 93(5): 1799-1812.

Bound, John and Gary Solon. 1999. "Double Trouble: On the Value of Twin-Based Estimation of the Return to Schooling." *Economics of Education Review* 18, 169-182.

Card, David. 1999. "The Causal Effect of Education on Earnings,” chapter 30 in The Handbook of Labor Economics, Volume 3A, 1999.

Card, David. 2001. “Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems.” *Econometrica* 69(5):127-1160.

Chou, Shin-Yi, Jin-Tan Liu, Michael Grossman, Theodore Joyce. 2007. "Parental Education and Child Health: Evidence from a Natural Experiment in Taiwan." NBER working paper No. 13466.

Cutler, David, and Adriana Lleras-Muney. (2008). "Education and Health: Evaluating Theories and Evidence." In: The Effects of Social and Economic Policy on Health, Editors J. House, R. Schoeni, G. Kaplan, and H. Pollack, forthcoming, Russell Sage Press.

Currie, Janet, and Enrico Moretti. 2003. "Mother’s Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings and Longitudinal Data." *Quarterly Journal of Economics* 118(4): 1495-1532.

de Walque, Damien. 2007. "Does Education Affect Smoking Behavior? Evidence Using the Vietnam Draft as an Instrument for College Education." *Journal of Health Economics* 26(5): 877-895.

Fuchs, Victor R. 1982. "Time Preference and Health: An Exploratory Study", in: Fuchs, V.R., (ed.), Economic Aspects of Health, University of Chicago Press, Chicago, pp. 93-120.

Griliches, Zvi. 1979. "Sibling Models and Data in Economics: Beginnings of a Survey." *Journal of Political Economy* 87(5): 37–64.

Grimard, Franque, and Daniel Parent. 2007. "Education and Smoking: Were Vietnam Draft Avoiders also More Likely to Avoid Smoking?" *Journal of Health Economics* 26(5): 896-926.

Groot, Wim. 2000. "Adaption and Scale of References Bias in Self-Assessments of Quality of Life." *Journal of Health Economics* 19(3): 403-420.

Grossman, Michael. 1972. "On the Concept of Health Capital and the Demand for Health." *Journal of Political Economy* 80(2): 223-55.

Grossman, Michael. 2004. "The Demand for Health, 30 Years Later: A Very Personal Retrospective and Prospective Reflection." *Journal of Health Economics* 23(4): 629-636.

Grossman, Michael. 2006. "Education and Non-market Outcomes." In Handbook of the Economics of Education Vol. 1, edited by E. Hanushek and F. Welch. Amsterdam: North-Holland, Elsevier Science, 577-633.

Hungerford, Thomas and Gary Solon. 1987. "Sheepskin Effcts in the Return to Education". *Review of Economics and Statistics* 69(1): 175-177.

Idler, Ellen I, and Yael Benyamini. 1997. "Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies." *Journal of Health and Social Behavior* 38(1): 21-37.

Isacsson, Gunnar. 1999. "Estimates of the Return to Schooling in Sweden From a Large Sample of Twins." *Labour Economics* 6(4): 471-489.

Jaeger, David A. 1997. "Reconciling the Old and New Census Bureau Education Questions: Recommendations for Researchers." *Journal of Business and Economic Statistics* 15(3), 300-309.

Kenkel, Donald. 1991. "Health Behavior, Health Knowledge, and Schooling." *Journal of Political Economy* 99(2): 287-305.

Kenkel, Donald, Dean Lillard, and Alan Mathios. 2006. "The Roles of High School Completion and GED Receipt in Smoking and Obesity." *Journal of Labor Economics* 24(3): 635-660.

Lahema, Eero, Pekka Martikainen, Mikko Laaksonen, and Akseli Aittomäki. 2004. "Pathways Between Socioeconomic Determinants of Health." *Journal of Epidemiology and Community Health* 58(4): 327-332

Leigh, Paul J., and Rachna Dhir. 1997. "Schooling and Frailty Among Seniors." *Economics of Education Review* 16(1): 45-57.

Lleras-Muney, Adriana. 2005. "The Relationship Between Education and Adult Mortality in the U.S." *Review of Economic Studies* 72(1): 189-221.

Lundborg, Petter, and Björn Lindgren. 2004. "Do They Know What They are Doing? Risk Perceptions and Smoking Behaviour among Swedish Teenagers." *Journal of Risk and Uncertainty* 28(3): 261–286.

Lundborg, Petter. 2007. "Smoking, Information Sources, and Risk Perceptions. New Evidence on Swedish Data." *Journal of Risk and Uncertainty* 34(3): 217-240.

Mazumder, Bhashkar. 2007. "How Did Schooling Laws Improve Long-Term Health and Lower Mortality?" Federal Reserve Bank of Chicago, WP 2006-23.

Miller, Paul, Charles Mulvey, and Nick Martin. 2005. "Birth Weight and Schooling and Earnings: Estimates from a Sample of Twins." *Economics Letters* 86(3): 387–392.

Nichols, R C, and Bilbro, W C. 1966. "The Diagnosis of Twin Zygosity." *Acta Genet.* 16(3): 265-275.

Oreopoulos, Philip. 2006."Estimating Average and Local Average Treatment Effects of Education when Compulsory School Laws Really Matter." *American Economic Review* 96(1): 152-175.

Rosenzweig, Mark R. and T Paul Schultz. 1982. "The Behavior of Mothers as Inputs to Child Health: the Determinants of Birth Weight, Gestation, and Rate of Fetal Growth", in: Victor R. Fuchs, ed., Economic Aspects of Health (University of Chicago Press for the National Bureau of Economic Research, Chicago) pp. 53-92.

Sander, William. 1995. "Schooling and Quitting Smoking." *Review of Economics and Statistics* 77(1), 191-199.

Spasojevic, Jasmina. 2003. "Effects of Education on Adult Health in Sweden: Results from a Natural Experiment." Ph.D. Dissertation, City University of New York Graduate Center.

Viscusi, W. Kip. 1990. "Do Smokers Underestimate Risks?" *Journal of Political Economy* 98(6): 1253– 1268.

Viscusi, W. Kip. 1991. "Age Variations in Risk Perceptions and Smoking Decisions." *The Review of Economics and Statistics* 73(4): 577–588.

Wolfe Barbara L, and Robert H. Haveman. 2002. "Social and Nonmarket Benefits from Education in an Advanced Economy", in: Yolanda K. Kodrzyki, ed., Education in the 21st Century: Meeting the Challenges of a Changing World (Conference Series 47, Federal Reserve Bank of Boston) pp. 97-131.

8 Tables

Table 1: Descriptive statistics.

Variables	**Means (std. err.)**		
	Main sample	**Twin sample**	**CPS**
Socio-economic and demographic			
Female (percent)	0.505	0.527	0.517
	(.009)	(.019)	
Age 25-34	0.215	0.226	0.276
	(.007)	(.016)	
Age 35-44	0.246	0.317	0.270
	(.007)	(.018)	
Age 45-54	0.236	0.249	0.192
	(.007)	(.016)	
Age 55-64	0.192	0.141	0.139
	(.007)	(.013)	
Age 65-74	0.109	0.066	0.122
	(.005)	(.009)	
White	0.880	0.934	0.848
	(.006)	(.010)	
Married or cohabitating	0.678	0.776	-
	(0.008)	(.015)	
High school graduate	0.297	0.318	0.342
	(.008)	(.018)	
Some college (no bachelor)	0.309	0.330	0.276
	(.008)	(.018)	
College graduate	0.283	0.291	0.226
	(.008)	(.017)	
Income ($1,000)	23,549	25,979	
	(463.724)	(1001.463)	
Health variables			
Health 0-10 scale	7.348	7.856	
	(.031)	(.057)	
Number of chronic conditions	2.569	1.982	
	(.049)	(.086)	
Smoking	0.243	0.213	
	(.007)	(.016)	
Physical activity	5.158	5.822	
	(.100)	(.214)	
Body Mass Index	26.893	26.025	
	(.104)	(.191)	
Overweight	0.495	0.490	
	(.009)	(.019)	
Work impact on health	2.590	2.612	
	(.023)	(.048)	
Work accident	0.610	0.545	
	(.064)	(.081)	
Health insurance	0.886	0.901	
	(.006)	(.012)	

Table 2: Regressions on self-reported health and the number of chronic conditions.

Variables	Main	Pooled	FE	Main	Pooled	FE
	Self-reported health			Number of chronic conditions		
Age	-0.038**	-0.051		0.100***	0.090*	
	(0.018)	(0.036)		(0.028)	(0.052)	
Age squared	0.000**	0.001*		-0.001**	-0.001	
	(0.000)	(0.000)		(0.000)	(0.001)	
Female	0.024	0.119		0.567***	0.695***	
	(0.063)	(0.115)		(0.098)	(0.168)	
White	-0.149	-0.588**		0.043	-0.661**	
	(0.096)	(0.230)		(0.151)	(0.335)	
Married/Part.	0.007	0.038	0.075	-0.090	-0.132	-0.275
	(0.071)	(0.144)	(0.198)	(0.112)	(0.210)	(0.261)
Income	0.003***	0.002**	0.003**	-0.003***	-0.002	-0.002
	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
High school	0.480***	0.569**	1.043**	-0.784***	-1.703***	-1.480***
	(0.114)	(0.254)	(0.422)	(0.180)	(0.370)	(0.556)
Some college	0.382***	0.658**	1.242***	-0.745***	-2.055***	-1.203**
	(0.114)	(0.255)	(0.440)	(0.180)	(0.372)	(0.579)
College degree	0.590***	0.904***	1.221**	-1.203***	-2.005***	-0.941
	(0.118)	(0.260)	(0.507)	(0.186)	(0.379)	(0.667)
Constant	7.672***	8.340***	6.476***	0.313	1.917	3.508***
	(0.422)	(0.874)	(0.435)	(0.664)	(1.274)	(0.572)
n	2877	642	624	2886	641	622

Table 3: Regressions on smoking.

Variables	Pooled	FE	Pooled	FE
	Smoking		Smoking at 16	
Age	0.009		-0.005	
	(0.010)		(0.012)	
Age squared	-0.000		0.000	
	(0.000)		(0.000)	
Female	0.032		-0.135***	
	(0.031)		(0.040)	
White	0.060		0.165**	
	(0.063)		(0.080)	
Married/Part.	-0.055	0.051	-0.052	-0.045
	(0.039)	(0.043)	(0.050)	(0.056)
Income	-0.001*	-0.000	0.001	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
High school	-0.183***	-0.070	-0.077	0.010
	(0.069)	(0.098)	(0.088)	(0.126)
Some college	-0.185***	-0.047	-0.121	-0.173
	(0.070)	(0.101)	(0.088)	(0.131)
College degree	-0.357***	-0.167	-0.226**	-0.023
	(0.071)	(0.116)	(0.090)	(0.150)
Constant	0.325	0.270***	0.728**	0.603***
	(0.238)	(0.100)	(0.303)	(0.130)
n	642	690	638	686

Table 4: Regressions on physical activity, BMI, and overweight.

Variables	Pooled	FE	Pooled	FE	Pooled	FE
	Physical activity		BMI		Overweight	
Age	-0.140		0.189*		0.018	
	(0.123)		(0.114)		(0.012)	
Age squared	0.001		-0.002		-0.000	
	(0.001)		(0.001)		(0.000)	
Female	0.773*		-1.526***		-0.230***	
	(0.401)		(0.371)		(0.039)	
White	1.053		-1.663**		-0.128	
	(0.796)		(0.744)		(0.078)	
Married/Part.	0.438	-0.228	-0.230	-0.083	0.009	-0.019
	(0.499)	(0.743)	(0.464)	(0.405)	(0.049)	(0.056)
Income	0.006	0.007	-0.004	-0.001	-0.000	-0.000
	(0.004)	(0.006)	(0.004)	(0.003)	(0.000)	(0.000)
High school	1.427	3.049*	-1.607*	0.751	-0.156*	-0.003
	(0.880)	(1.574)	(0.825)	(0.839)	(0.087)	(0.127)
Some college	1.827**	3.369**	-2.266***	0.027	-0.160*	-0.009
	(0.884)	(1.642)	(0.829)	(0.881)	(0.087)	(0.132)
College degree	2.198**	3.314*	-3.160***	0.019	-0.200**	0.050
	(0.899)	(1.890)	(0.840)	(1.018)	(0.089)	(0.151)
Constant	9.252***	5.909***	26.406***	25.757***	0.510*	0.519***
	(3.030)	(1.623)	(2.814)	(0.871)	(0.298)	(0.131)
n	639	646	618	625	642	690

Table 5: Regressions on occupational hazards.

Variables	Pooled	FE	Pooled	FE
	Job affects health		Number of work accidents	
Age	-0.004		0.028	
	(0.036)		(0.061)	
Age squared	-0.000		-0.000	
	(0.000)		(0.001)	
Female	-0.195**		-0.335**	
	(0.098)		(0.169)	
White	-0.279		0.365	
	(0.203)		(0.347)	
Married/Part.	-0.131	-0.053	0.087	0.538
	(0.121)	(0.201)	(0.207)	(0.385)
Income	0.000	0.001	-0.002	0.000
	(0.001)	(0.002)	(0.002)	(0.003)
High school	-0.193	0.144	-0.134	-0.447
	(0.237)	(0.509)	(0.407)	(0.974)
Some college	-0.359	-0.107	0.277	0.201
	(0.237)	(0.507)	(0.406)	(0.971)
College degree	-0.305	0.010	-0.124	-0.981
	(0.237)	(0.560)	(0.406)	(1.073)
Constant	3.748***	2.561***	0.222	0.468
	(0.805)	(0.510)	(1.385)	(0.974)
n	499	481	498	482

Table 6: Regressions on health insurance.

Variables	Pooled	FE
	Health insurance	
Age	0.003	
	(0.007)	
Age squared	-0.000	
	(0.000)	
Female	0.003	
	(0.023)	
White	0.057	
	(0.046)	
Married/Part.	0.053*	-0.011
	(0.029)	(0.044)
Income	0.089*	0.000
	(0.052)	(0.000)
Health insur.		
High school	0.133**	-0.014
	(0.052)	(0.095)
Some college	0.179***	-0.119
	(0.053)	(0.098)
College degree	0.001**	-0.036
	(0.000)	(0.112)
Constant	0.538***	0.925***
	(0.176)	(0.096)
n	638	614

Table 7: Pooled twin sample and twin FE of the health returns to education by parents' education.

Variables	Pooled	FE	Pooled	FE
	Self-reported health		Chronic conditions	
High school	1.911**	1.251***	-1.768	-1.289**
	(0.741)	(0.441)	(1.086)	(0.581)
Some college	1.699**	1.244***	-0.904	-0.985
	(0.744)	(0.459)	(1.090)	(0.604)
College degree	1.711**	1.250**	-1.341	-0.933
	(0.747)	(0.514)	(1.095)	(0.677)
High school * parents' educ	-0.882*	-0.382**	-0.076	0.093
	(0.497)	(0.174)	(0.729)	(0.229)
Some college * parents' educ	-0.702	-0.134	-0.737	0.059
	(0.487)	(0.127)	(0.714)	(0.168)
College degree * parents' educ	-0.600	-0.174	-0.470	0.227
	(0.481)	(0.108)	(0.706)	(0.143)
Parent's educ	0.605		0.508	
	(0.468)		(0.686)	
n	609	294	608	

Table 8: Differences within twin pairs in early life.

Variables	Parents emphasized differences between the twins	Twins shared same classroom	Twins shared playmates
Always	1.5%	35.4%	53.1%
Most of the time	6.2%	21.8%	36.6%
Some of the time	6.8%	28.6%	8.6%
Never	85.5%	14.2%	1.8%

Table 9: Correlation of education and other characteristics between twin pairs and within twin pairs.

Variables	Correlation between average twin-pair education and average twin-pair characteristics	Correlation between within-twin-pair differences in education and within-twin-pair characteristics
	Education	ΔEducation
Birthweight	0.0001	0.0001
Phys. health at 16	0.0572	0.068
Ment. health at 16	0.0404	0.0446
Mother: time and attention	0.1568**	0.0049
Father: time and attention	-0.0860	0.0641
Mother: ove and affection	-0.0543	0.0201
Father: love and affection	-0.0038	0.1056**
Mother: strictness about rules	-0.0076	0.0521
Father: strictness about rules	0.1081*	-.0044
Mother: harsch when punishing	0.0369	0.0256
Father: harsch when punishing	0.1120*	0.0106
Mother: relationship rating	-0.0036	-0.0445
Father: relationship rating	-0.0268	-0.0075
Mother: expectations	-0.2266***	-0.0128
Father: expectations	-0.0778	-0.0044
Mother: physical abuse	0.1269**	0.0446
Father: physical abuse	0.2015***	0.0275