Adolescent Health and Adult Earnings: Evidence from a Large Sample of Siblings and Twins

Petter Lundborg[a]

Anton Nilsson[b]

Dan-Olof Rooth[c]

January 2012

[a]Department of Economics and Centre for Economic Demography at Lund University, VU University Amsterdam, Tinbergen Institute, Netspar, HEP and IZA. E-mail: Petter.Lundborg@nek.lu.se.

[b]Department of Economics at Lund University. E-mail: Anton.Nilsson@nek.lu.se.

[c]Linnaeus University, Centre for Economic Demography at Lund University, CReAM and IZA. E-mail: Dan-Olof.Rooth@lnu.se.

Abstract

We study the relationship between health status and adult earnings using a unique dataset that covers almost the entire population of Swedish males born between 1950 and 1970. The health information is obtained from medical examinations during the mandatory military enlistment tests at age 18, which we have further linked to register data on adult earnings. We find that most types of major conditions have long-run effects on future earnings with the largest effects resulting from mental ones. Including sibling fixed effects or twin-pair fixed effects reduces the magnitudes of the estimates, although remaining substantial.

JEL classification: I1, J24, J31

Keywords: earnings, health, specific conditions, siblings, twins

1. Introduction

An adult's success on the labor market, to a large extent, depends on his childhood or adolescent health status. This is suggested by a growing literature that links measures of health status in early stages of life to adult outcomes, such as education and earnings (e.g., Currie and Hyson 1999; Currie et al. 2010; Smith 2009). Since children with poor health are disproportionally to be found in families of lower socioeconomic status, this implies that health early in life may be an important factor in understanding how socioeconomic status is transmitted across generations. In order to design policies that prevent disadvantaged children from falling behind, and thus increase intergenerational mobility, more evidence on the impact of poor health on adult outcomes is clearly needed, however.

In this paper, we study the relationship between health status early in life and adult earnings, using a unique dataset that covers almost the entire population of Swedish males born between 1950 and 1970. The dataset includes information on health status obtained from a medical examination at the age of 18, conducted during the mandatory military enlistment test procedure in Sweden. In addition, the data includes cognitive and noncognitive test scores from the military enlistment records as was well as register-based data on adult earnings, occupation, educational attainment, family links, parental educational attainment, and parental earnings.

Our data gives us a number of advantages. First, it is based on obligatory assessments of the individual's health conducted by a physician. As a result, measurement errors originating from differences in health-seeking behaviors, justification bias or differences in expectations, which are typically present in sources like hospital and insurance records, sickness absence records or standard self-reports, are less of an issue. Moreover, the data contains a large number of different diagnoses, reflecting both mental and physical problems.

Second, as military enlistment was mandatory during the study period, our data covers more or less the entire population of males. This gives our results an unusually high degree of representativeness. The unusually large sample size also allows us to consider the impact of diagnoses that are somewhat rarer and thus hard to estimate in smaller samples.

Third, while previous research on the relationship between early health and adult labor market outcomes has mostly used measures of *very* early health, such as birth weight, few studies have considered the importance of health status after birth. At birth, many diseases have not yet developed, or may be difficult to measure. It is thus somewhat unclear to what extent birth weight and similar indicators of infant health are reasonable proxies of overall health status or wage earning capacity early in life. By considering health status *right before* the typical age of labor market entry, a much more comprehensive set of health problems can be included, yet reverse causality should not be a critical issue.

Fourth, our data allows us to compare outcomes and health between almost 275,000 sibling brothers. This is important, since various unobserved family-specific factors may affect both earnings and health, creating a spurious relationship between the two. By including sibling fixed effects, we are able control for such unobserved factors at the family level. In addition, we are able to identify a relatively large sample of monozygotic twins in our data. This allows us the further difference out the influence of unobserved endowments, such as genes. With the exception of studies using birth weight and other measures of health already around birth, we believe our study is the first to estimate the effect of early health on subsequent earnings using data on monozygotic twins.

Fifth and finally, our data includes cognitive and noncognitive test scores. We are thus able to control for ability differences between siblings, while evaluating the impact of childhood health. Without such controls, one risks confounding the effect of health with that of ability, since some previous studies, such as Case and Paxson (2008) and Lundborg et al.

(2010), show that the returns to height, when treated as an indicator of childhood health, is substantially reduced when accounting for cognitive skills.

As it turns out, our results suggest a strong relationship between health at age 18 and adult earnings both with and without sibling fixed effects, although the estimates are in virtually all cases reduced with the introduction of the fixed effects. In particular, having a mental condition at age 18 gives rise to an earnings penalty of about 20 percent when sibling fixed effects are included and of about 8 percent when controls for cognitive and noncognitive ability are also added. The latter estimate may be viewed as a lower bound of the total causal effect, however, as both cognitive and noncognitive ability may be partly determined by health status in young ages.

The rest of the paper is organized as follows. Section 2 provides a background discussion and reviews some relevant literature. In section 3 we introduce the empirical model and describe the data, the construction of the variables, and the sample restrictions. Section 4 presents the results, whereas section 5 provides some concluding remarks.

2. Background

Medical scientists have traditionally assumed a causal relationship from socioeconomic status to health. A growing body of research in both medicine and economics instead points to relations in the opposite direction. In particular, evidence is accumulating on the importance of health early in life for a range of different outcomes. This literature has mostly considered measures of *very* early life health. For instance, a number of studies in the economics literature have shown the importance of birth weight for outcomes such as education and earnings (e.g., Behrman and Rosenzweig, 2004; Black et al., 2007; Currie and Hyson, 1999; Royer 2009).

Recently, a few studies have made use of other measures of childhood or adolescence health. Using data from the 1958 British cohort study, Case et al. (2005) showed that adults who had suffered from chronic conditions during childhood (at age 7 or 16) had lower educational attainment, earnings, and employment probabilities compared to other adults.

Smith (2009) used a retrospective measure of self-assessed health up to age 16 from the Panel Study of Income Dynamics (PSID) and contrasted outcomes between siblings. The results suggested that poor self-reported health at age 16 had a significant and negative relation with later earnings and other economic outcomes. In fact, the sibling fixed effects estimates came out larger than the corresponding OLS estimates, which may possibly be explained by the fact that family-specific measurement errors got differences out in the fixed effects models.

In Currie et al. (2010), data from public health insurance records for 50,000 children and adolescents aged 18 and younger born in the Canadian province of Manitoba was used to estimate the relation between childhood health and later outcomes. The results suggested significant and negative effects of both physical and mental health problems on various later outcomes, such as school grades and the probability of receiving social assistance, in models with sibling fixed effects.[1]

Salm and Schunk (2011) consider the effect of childhood health on cognitive and verbal development at age six in a sample of German children. Using sibling fixed effects models, their results suggest that at least mental health problems can account for a substantial fraction

[1] In addition to these studies using individual-level measures of health status early in life, research has considered the long-run effects of health shocks affecting certain cohorts or geographical populations. For example, Almond (2006) finds that individuals who *in utero* were exposed to the 1918 influenza panepidemic displayed increased rates of physical disability, lower educational attainment, lower socioeconomic status in terms of occupational choice, and lower earnings.

of the variation in cognitive and verbal test scores. Similar evidence is documented by Currie and Stabile (2007) who find that in especially ADHD has large negative effects on future test scores and schooling attainment. These results are important, because they point to one mechanism whereby early life health problems may affect later socioeconomic achievement.

3. Data and method

3.1 Data

Our dataset is based on two main sources that have been linked together. First, information on individuals' health status, cognitive ability and noncognitive ability has been obtained from military enlistment records from 1969-1997. These are provided by the Swedish National Service Administration ("Pliktverket"). Second, information on educational attainment, occupation and earnings is provided by Statistics Sweden ("Statistiska centralbyrån"). Most individuals can be linked to their parents, which enables us to control for environmental effects at the family level both using parental characteristics and using sibling fixed effects.

At the time under study, the Swedish military enlistment test was mandatory for men, with exemptions only granted for institutionalized individuals, prisoners, and individuals living abroad. Individuals usually took the test at the age of 18 or 19.[2] The military enlistment records include up to six medical conditions for every individual, using the ICD8 classification (WHO, 1967). There is also a unidimensional global health measure, which is based on the severeness of the individual's health conditions (both physical and mental) and is used to determine his suitability with respect to type of military service.

[2] Up until 1972, a majority of individuals were called to undergo the military enlistment test the year they turned 19, whereas in later years most individuals underwent the enlistment test the year they turned 18.

All diagnoses are recorded by a physician and are based on a health declaration form that the individual has filled in at home and has to bring with him, combined with a general assessment of the individual's health lasting for about 20 minutes. The individual is expected to bring any doctor's certificate, health record, drug prescription or similar proving that he actually suffers from the conditions he has reported in his health declaration, making "cheating" difficult. If the individual does not provide such evidence, he may be forced to visit a hospital or health care center to obtain a doctor's certificate, and then continue the military enlistment procedure at a later point.

It is a significant advantage of our data that health is not simply self-reported, but also based on obligatory assessments. As a consequence, measurement errors for example originating from differences in health-seeking behaviors or in health awareness, which may be present in sources like hospital and insurance records or standard self-evaluations, should be less of an issue. One potential disadvantage is that conditions that are too mild to affect the individual's suitability for military service have not have been recorded, although it seems quite likely that these will also be the conditions that are also too mild to affect individual success on the labor market.

The global health measure included in our data is expressed with letters from A to M (except "I"), or "U", "Y", or "Z". The closer to the start of the alphabet the letter assigned to the individual is, the better his general health status is considered to be. "A" thus represents more or less perfect health, which is necessary for "high mobility positions" (for example light infantry or pilot) and has been assigned to about two-thirds of all individuals for which there is non-missing data. For combat positions, individuals must have been assigned at least a "D"; individuals with a "G" or lower are only allowed to function in "shielded positions" (including for example meteorology or shoe repairing). Individuals assigned a "Y" or "Z" (in total 6 % of all individuals in our "full sample", as defined below) are not allowed to undergo education within the military. "U" indicates that global health status has not been decided, and we treat this as missing. The global health measure to be used in our analysis is created by transforming "A" into 0, "B" into 1 etc., "Y" into 12 and "Z" into 13.

We use 13 broad categories of health conditions, which in total cover almost 97 percent of all conditions reported in our "full sample". In creating these categories of diagnoses, we followed the "chapters" provided by the ICD8 classification, with the exception that we found it natural to split the category of "conditions of the nervous system and sensory organs" into one category of conditions of the nervous system[3] and one category of conditions of the sensory organs[4]. This leaves us with the following categories of health conditions: tumors, endocrine, nutritional and metabolic conditions, mental conditions, conditions of the nervous system, conditions of the sensory organs, circulatory conditions, respiratory conditions, digestive conditions, genito-urinary conditions, skin conditions, conditions of musculoskeletal and connective tissue, congenital anomalies, and injuries and poisonings.

In addition to these broad categories of health problems, we analyze a number of specific types of diagnoses. These include diabetes (an endocrine condition), neurosis, personality disorder and alcoholism and drug dependence (mental conditions), epilepsy and migraine (conditions of the nervous system), asthma and hay fever (respiratory conditions), and vertebrogenic pain syndrome (a condition of musculoskeletal and connective tissue).[5] Previous literature has documented relations between all these conditions and adverse labor market outcomes, however in most cases only simultaneous relationships using sources like

[3] ICD8: 320-358.

[4] ICD8: 360-389.

[5] The ICD8 codes for these are 250, 300, 301, 303-304, 345, 346, 493, 507, and 728, respectively.

sickness absence records or standard self-reports.[6] Asthma and mental problems have received particular interest in the recent early health literature (e.g., Currie et al., 2010; Salm and Schunk, 2011).

For the study period, cognitive ability was measured using four subtests (logical, verbal, spatial, and technical), each of which was graded on a scale between 0 and 40. We transform these test scores to a single variable as follows. First, test scores are normalized by birth year using all individuals for which there are non-missing test scores on the subtest in question.[7] For every individual, we then calculate the average of his non-missing normalized test scores. Finally, this variable is normalized using all individuals in the dataset.

Noncognitive ability is measured on a scale between 1 and 9, which approximates a normal distribution. The assignment of this number is done by a psychologist, based on a semi-structured interview lasting for about 25 minutes. The objective of this interview is to determine the individual's ability to cope with the psychological requirements of the military, and in particular this implies an assessment of personal characteristics such as willingness to assume responsibility, independence, outgoing character, persistence, emotional stability and power of initiative (Lindqvist and Vestman, 2009). In addition, an important objective of the interview is to identify individuals who are considered particularly unsuited for military service, which includes individuals with antisocial personality disorders, individuals with difficulty accepting authority, individuals with difficulties adjusting to new environments and individuals with violent and aggressive behavior (Andersson and Carlstedt, 2003; Lindqvist and Vestman, 2009). It is thus by definition true that this variable is endogenous at least to some mental conditions. For our empirical analysis, we proceed similarly as for cognitive

[6] See, for example, Ng et al., (2001), Ettner et al. (1997), Ettner et al. (2011), Chatterji et al. (2007), Famulari (1992), Ferrari (1998), Krahn et al. (1996), Malone et al. (1997), and Fanning (1981).

[7] We treat zeros as missing since individuals should get at least a few points by pure chance.

ability and standardize the noncognitive ability variable by birth year using all individuals with non-missing data on this variable**.** Lindqvist and Vestman (2009) have previously shown that both our cognitive and noncognitive measures strongly predict adult earnings in the population of Swedish males.

We use data on individuals' own educational attainment, occupational choice and annual earnings from 2003. The latter refers to income from work and self-employed income. Educational attainment is expressed in terms of the highest degree attained. Our measure of years of schooling is assigned based on the standard number of years of schooling associated with this degree. We choose to standardize this variable in order to make it comparable with cognitive and noncognitive ability. There is also information on parents' earnings and educational attainment, where the former is measured in 1980 and the latter in 1999.

Occupation is categorized into 115 different groups according to SSYK (Standard for Swedish Occupational Classification), a three digit occupational classification code similar to the international classification (ISCO). These indicators are then entered as fixed effects in (some of) the regressions. Doing this we can examine whether poor health affects earnings though occupational choice.

3.2 Sample constructions

The oldest cohort available is born in 1950. At the younger end, we choose to only include individuals born up until 1970. This is for two reasons. First, younger individuals may not have finished their education as of 2003 or may be temporarily employed in jobs that do not fully correspond to their productivity level. Second, beginning in 1989, diagnoses were reported by the National Service Administration using the ICD9 instead of the ICD8 classification. Since ICD8 and ICD9 codes are not completely comparable, it would thus make our results somewhat less straightforward to interpret had we included later born individuals.

In order to avoid selection issues, we exclude the small number of women volunteering for the military. Similarly, we only include native Swedes, that is, individuals born in Sweden to Swedish-born parents. This is to sort out non-Swedish citizens, who are not obliged to undergo military service in Sweden and, in addition, avoid issues of ethnic discrimination.

Our sources provide us with a population of 951,017 native Swedish males born between 1950 and 1970. We construct our "full sample" as follows. First, excluding individuals who lack data on (global) health status (18.6 %), our sample includes 774,529 individuals.[8] Second, individuals not reported to have positive annual earnings in 2003 are excluded (8.3 %), giving us a sample of 710,018 individuals.

In addition to this "full sample", we make use of a sibling subsample. This subsample includes 273,296 individuals. Applying sibling fixed effects on these individuals allows us to control for unobserved characteristics at the family level.

Furthermore, in an attempt to even more fully account for unobserved characteristics, we have also obtained data on twins and their zygosity from the Swedish Twin Registry[9] ("Svenska tvillingregistret") and applied twin-fixed effects. This data on twins includes 6,867 individuals born between 1950 and 1970, of which 2,316 are monozygotic (MZ), 3,185 are dizygotic (DZ) and 1,366 are of unknown zygosity. In particular, MZ twins are of interest because these share the same DNA, allowing us to fully control for genetic factors.

[8] For apparently random reasons, data is missing for certain time periods. For example, we have no data on individuals taking the enlistment tests in 1985. About 3 % of the population lack data because they didn't undergo the military enlistment procedure.

[9] Detailed information on the Swedish Twin Registry can be found in Lichtenstein et al. (2006).

Some individuals have missing values on years of schooling, cognitive ability, noncognitive ability or parental characteristics. In our OLS regressions, we used sample averages and created a binary variable taking on the value one when there is missing information on a variable. Similarly, for our sibling or twin-fixed effects estimates, we make within-family differences equal zero by setting the missing values equal to family averages and add a binary variable indicating the presence of missing data.

Descriptive statistics for our full sample and for our sibling subsample is shown in Table 1 and 2. Regarding health, it can be seen that about 50 percent of all individuals are diagnosed with some condition in both samples. Conditions of the sensory organs and conditions of musculoskeletal and connective tissue are the largest groups, followed by respiratory conditions.

3.3. Empirical method

For the full sample, we run regressions of the following form:

$$1) \quad Log\ Earnings_i = a + b*H_i + c*X_i + d*F_i + e_i,$$

where i is index for individual. H is either a scalar representing global health, a vector of the different categories of health conditions (including an indicator for “other categories”), or a vector of the specific conditions that we analyze (including an indicator for “all other”). X is a vector of individual characteristics and F is a vector of parental characteristics. As ability and other productivity-related characteristics to some extent are expected to be determined by health status, it is not fully clear what to control for, given our interest in the causal effects of health. We therefore run a number of models including different sets of control variables. Given that correlations *partially* reflect causality, that all controls included have the effect of lowering the estimated health penalty, and that no other confounding mechanisms are at play, estimates from models without these controls can be seen as upper bounds whereas estimates including all of these can be seen as lower bounds of the effects of health status at the age of 18 on earnings later in life.

In addition to Equation 1, we run sibling (or twin) fixed effects regressions of the following form:

2) $$Log\ Earnings_{ij} = a + b{*}H_{ij} + c{*}X_{ij} + d{*}f_{i} + e_{ij},$$

where j is index for family. In this equation, f_j represents sibling fixed effects capturing characteristics that are common to the brothers and affect earnings capacity. Again, X is a vector of individual characteristics. The identification of b relies upon within-siblings variation in health variables.[10]

4. Results

4.1 Global health – the full sample

As can be seen in Model A in Table 3, global health status at the age of 18 is a strong predictor of adult earnings. According to the estimate, an increase in bad health of ten steps, which corresponds to the difference between "perfect" health and a health status just slightly better than required for military service, implies a 25 percent decrease in earnings. Given

[10] Sibling fixed effects may be inappropriate if the health status of an individual affects the outcomes of his siblings via, for example, social interaction or parental treatment. Given that an individual's health problems have mostly negative spillovers on his siblings, estimates are downward biased. Then again, coefficients from specifications including full controls (including the sibling fixed effects) can be viewed as lower bounds.

conventional OLS estimates of the returns to schooling, this is an effect equivalent of about 3-5 years of education. As no controls have been included, however, this estimate may be viewed as an upper bound.

Controlling for parental characteristics (i.e., father's earnings, father's years of schooling, and mother's years of schooling) leaves the health penalty almost unchanged at 23 percent (Model B). Having a cognitive or noncognitive ability level one standard deviation above the mean is associated with having about 10 percent higher earnings, and the inclusion of these have strong effects on the health penalty. Controlling for cognitive ability reduces the coefficient on global health by a fourth, and adding the control for noncognitive ability further reduces the coefficient by a half to 8.7 percent.

While schooling is strongly related to earnings, the inclusion of this variable has a rather small effect on the health penalty, as can be seen when comparing Model F to Model B. There is thus little evidence in our data of schooling as an important mediator of the health-earnings relationship. In addition, the inclusion of schooling eliminates the previously significant impacts of parents' educational attainment.

Model G and H include occupation-fixed effects. Without the controls for parental characteristics, ability or educational attainment, an increase in bad health of ten steps corresponds to 15 percent lower earnings within the same occupation. With controls for all variables available (parental characteristics, cognitive skill, noncognitive skill, years of schooling, and occupation), the health penalty within a given occupation reduces by a half to 7.5 percent, which is not much smaller than the coefficient obtained with full controls for ability and no occupation-specific effects. Our findings suggest that individuals in poor health are sorted into less well paid jobs, but that neither this nor lower educational attainment of individuals in poor health is the major explanation for the negative relationship between health and earnings.[11]

4.2 Global health – the sibling subsample

In order to fully control for unobserved family-level characteristics, we now restrict attention to the 273,296 brothers in the sample. We estimate models with the same individual-specific controls as for our full sample, enabling comparison with our previous results. The results are displayed in Table 4.

In the first specification (Model A), we estimate the relationship without any controls and without sibling fixed effects. As can be seen, this produces an estimate that is very close to the corresponding estimate in table 3, indicating that the subsample of brothers is quite representative of a larger population. When applying sibling fixed effects (Model B), the estimate is reduced by a third to -17.5 percent. This is a large reduction compared to Model B in Table 3, where father's earnings, father's years of schooling and mother's years of schooling were controlled for, suggesting that these measures are rather imperfect proxies of family background.[12]

Including cognitive or noncognitive ability again substantially lowers the health penalty (Models C-E). In particular, after including noncognitive ability, estimates become very close

---

[11] Using Swedish military enlistment data, Lundborg et al. (2011) find no evidence of a relationship between health at age 18 and educational attainment within pairs of MZ twins. Given that MZ twins are representative of a larger population, this may suggest that health does not causally affect educational attainment.

[12] There is no evidence that the comparatively small reduction of the health penalty that occurs when introducing parental characteristics in Table 3 is due to the high prevalence of missing values on parental characteristics. In fact, the health penalty reduces by a somewhat *smaller* amount (to 0.235) when excluding individuals with missing data on one or several parental characteristics.

to the ones obtained in Table 3, showing that unobserved characteristics at the family level are no longer important. Furthermore, including years of schooling (Model F) has little effect, again speaking against the possibility of schooling being an important mediator in the relationship between early health and later labor market success.

As before, adding occupation-specific effects has a limited effect on the bad health penalty; it is reduced to 13.1 percent without controls for cognitive ability, noncognitive ability and schooling, and to 7.7 percent when these controls are included. The main explanation for the health-earnings relationship is not that individuals in poor health work in other occupations than healthy people, but instead that individuals in poor health earn lower incomes in the same occupations.

4.3 Global health – twins

Our results for twins, using global health and no controls, are displayed in Table 5. We separately run the analysis on all twins, on DZ (dizygotic) twins only, and on MZ (monozygotic) twins only. While estimates based on these subsamples are uncertain due to the limited number of observations, a number of important results show up.

First of all, the OLS regressions (Model A, C and E) indicate that these subsamples are representative of a larger population. Moreover, running specifications including twin-fixed effects (Model B, D and F), estimates are reduced in similar magnitudes as previously observed for siblings. This shows that non-twin siblings may be quite enough to draw relevant conclusions regarding the health-earnings relationship.

We have also run twin-fixed effects models including controls for cognitive and/or noncognitive ability (not reported). In these specifications, the health penalty became statistically insignificant among MZ twins, and also among twins in general when noncognitive ability was included. These results imply that we cannot fully reject the possibility that health has no direct effect on earnings. Noncognitive ability was significant in all these specifications (at the 5 percent level), however, which means that there is at least statistical evidence of an indirect effect of health on earnings, given that noncognitive ability is partly determined by health.

4.4 Categories of health conditions

Regression results for specific categories of health diagnoses are shown in Table 6-7. In Table 7, sibling fixed effects are used. As before, the OLS regressions indicate that siblings are quite representative of the larger sample.

As can be seen, most types of conditions are strongly statistically significantly associated with lower earnings, irrespective of the specification. Tumors, circulatory conditions and skin conditions are the noteworthy exceptions, as these in most cases turn out insignificant, although one may argue that the inability to obtain significant effects of tumors in particular might be due to the small number of observations; only 412 individuals in the sibling sample are diagnosed with a tumor. At the same time, the coefficient estimates that we obtain for these conditions, including the significant ones, are small in comparison to other coefficients, so there is little indication of important causal effects of tumors, circulatory conditions and skin conditions on adult earnings. These are quite interesting findings because both circulatory conditions and cancer are the two most common causes of mortality in Sweden as well as in other industrialized countries (Socialstyrelsen, 2008; WHO, 2004), striking in particular individuals of lower socioeconomic status (e.g., Claussen et al., 2003; Ward et al., 2004). Our findings provide no evidence that the relationship between these conditions and socioeconomic achievement may be bidirectional.

The inclusion of sibling fixed effects in general has the effect of reducing the health penalties in magnitude (Model B in Table 7 compared to Model A). Among the significant diagnoses, the only exception is congenital anomalies, suggesting that individuals with congenital anomalies come from relatively advantaged family backgrounds. For respiratory conditions there is also a similar story, as this one is insignificant and close to zero in the sibling subsample without sibling fixed effects and even significantly positive in some specifications in the full sample, but becomes significantly negative when applying the sibling fixed effects.[13]

In addition to specific categories of diagnoses, these tables report the effect of being diagnosed with *any* health condition. Without controls, being diagnosed with any condition is estimated to reduce earnings by about 12 percent. When including sibling fixed effects, this coefficient is reduced to a negative of 7 percent, and controlling for cognitive and noncognitive ability further reduces it to minus 3 percent. Again, the health penalty is to the largest part driven by individuals in poor health earnings lower incomes within the same occupations compared to healthy people. This pattern also emerges for all significant categories of health conditions, and for all specific health diagnoses with significant effects (see below).

Some categories of health conditions have very large effects. In particular, endocrine/nutritional/metabolic conditions, mental conditions and conditions of the nervous system have coefficients similar to or even higher than the ones obtained for schooling,

[13] This finding regarding respiratory conditions is probably due to the fact children to smokers are more likely to get respiratory problems (e.g., Cook and Strachan, 1999), and that smoking in Sweden during this period of time was more common among among individuals with higher levels of educational attainment. The latter is rather well-known, but there seems to be no study written on the topic. Using the Swedish Level of Living Survey from 1968, which was based on a random sample of the adult population this year and included 6,520 individuals, we regressed a smoking dummy on a dummy indicating "higher than basic education" and obtained a positive coefficient of 0.06; t = 4.88.

cognitive ability and noncognitive ability. We find that irrespective of the specification, mental disorders have the by far strongest impact on earnings of all classes of conditions. For example, according to Model B in Table 7, which controls for sibling fixed effects, individuals with mental conditions earn on average 20 percent less, and according to Model G which also includes occupation fixed effects, these individuals earn on average 15 percent less. As expected, the effect of mental conditions is reduced heavily (by about 50 percent) when controlling for noncognitive ability, and with full controls we find that mental conditions are associated with 7.4 percent lower earnings.

Regardless of the specification, endocrine, nutritional and metabolic diseases stands out as the second most severe diagnose category; it has a coefficient of -0.10 when sibling fixed effects are included (Model B in Table 7), -0.08 when occupation fixed effects are added (Model G), and -0.04 when cognitive ability, noncognitive ability and schooling are also included (Model H)). As number three, conditions of the nervous system turn up with coefficients between -0.09 when only sibling effects are included (Model B) and -0.04 with full controls (Model H).

Other categories of conditions have generally much smaller effects. For example, conditions of the sensory organs only have between 15 and 25 percent of the effect of mental conditions, depending on the specification. Injuries and congenital anomalies, which were for example examined by Currie et al. (2010), have similar effects. Still, these conditions imply between 4 and 5 percent earnings reductions in the model with sibling fixed effects (Model B in Table 7).

Congenital anomalies are of special interest because these by definition develop prior to birth, that is, long before schooling investment decisions, occupational choice, and before cognitive and other individual abilities have stabilized.[14] The argument for controlling for such variables is thus rather weak in the case of congenital anomalies, as these variables should primarily be viewed as intermediating factors. We conclude that the causal effect of congenital anomalies is likely to be in the neighborhood of 5 percent, although the direct effects (the effects not mediated by abilities, schooling and occupation) is only about half of this.

4.5 Specific diagnoses

Table 8-9 show our results for the narrow classes of diagnoses that we have selected, where Table 9 is the one including sibling fixed effects. As before, schooling has little effect on the estimates and the health penalty is mostly a within-occupation effect. According to all models, alcoholism and drug dependence is the health problem producing the most severe effects. This condition has a penalty of more than 40 percent when no controls are included (Model A in Table 8 and 9). The penalty falls to 26 percent when controls for sibling fixed effects are included (Model B in Table 9), and further drops to 11 percent when full controls are added. This effect is still twice as large as the effects of cognitive skill, noncognitive skill, or schooling (Model H).

Besides alcoholism and drug dependence, personality disorder and neurosis tend to have the largest coefficients. The negative effects of both these types of conditions amount to about 30 percent without any controls (Model A in Table 8 and Model A in Table 9), 20 percent when sibling fixed effects have been included (Model B in Table 9) and fall to about 7 percent when controls for ability, schooling and occupation are also added (Model H in Table 9). It can be seen in Table 9 that the coefficient on personality disorder always reduces by a

[14] See Cunha et al. (2006) for an extensive discussion of critical periods in human ability formation. IQ, for example, stabilizes around age 10, whereas noncognitive abilities may be malleable up until the early 20s.

larger percentage than the coefficient on neurosis and the coefficient on alcoholism and drug dependence when controls are added, indicating that either confounding characteristics related to personality disorders or indirect paths between personality disorders and earnings are comparatively important.

Regarding alcoholism and drug dependence, one should note that this condition is likely to develop quite late in adolescence, when IQ and other personal characteristics have mostly already stabilized. The argument for controlling for cognitive and noncognitive ability is thus quite strong in this case, tightening the upper bound of the estimate of the causal effect to about 18 percent.

In addition to these mental diagnoses, diabetes and epilepsy are found to have negative effects on earnings. Both these amount to about 15 percent without controls and are almost unaffected with the inclusion of sibling fixed effects, suggesting that family-level skills are not importantly related to these conditions. On the other hand, the coefficient of diabetes is strongly reduced in magnitude by the inclusion of noncognitive ability and even becomes insignificant in most specifications including this variable, possibly indicating that diabetes has important adverse effects of noncognitive ability.

Regarding other types of conditions, that is, migraine, asthma, hay fever and vertebrogenic pain syndrome, everything except asthma is statistically significant in most of the pooled models, although with comparatively small coefficients. When including sibling fixed effects, all these coefficients approach zero and become insignificant in virtually all cases, which means that there is little evidence of a health penalty of any of these conditions. These insignificant results cannot be blamed on too little variation in the data, as the confidence intervals are all quite tight around zero.

5. Conclusion

Using a unique dataset covering almost the entire population of Swedish males born between 1950 and 1970, this study establishes that health at the age of 18 is strongly related to labor market success in middle adulthood. Controlling for cognitive ability, noncognitive ability and unobserved characteristics at the family level lowers the estimates, but important effects remain. Compared to cognitive and noncognitive ability, the inclusion of schooling has less influence on the bad health penalty and we can conclude that neither occupational choice nor educational attainment acts as a main mediator in the health-earnings relationship.

In addition to siblings, we examined subsamples of twins in an attempt to additionally control for unobserved characteristics. While these estimates are more uncertain due to the limited number of observations, we can confirm that global health is negatively related to later earnings, at least in models without further controls. It is reassuring that the health penalty for global health is of about the same size in our models with twin-fixed effects as in our models with sibling fixed effects, since this lends credibility to our sibling-based estimates in general.

Due to data availability, previous research regarding the long-term effects of health status early in life has in general only used unidimensional health measures such as birth weight or retrospectively self-reported global health. As a major contribution of this paper, we conduct an analysis of a large number of broad categories of health conditions, as well as more narrow classes. This analysis shows that while most categories of diseases seem to have important long-run effects, the strongest ones are detected for mental conditions, with rather similar effects of the particular types of diagnoses examined within this group: neurosis, personality disorder and alcoholism and drug dependence. Above all, our findings suggest that reducing the prevalence of these conditions should be expected to have substantially positive effects on affected individuals' long-run productivity and material well-being.

References

Almond, Douglas (2006), "Is the 1918 Influenza Pandemic Over? Long-Term Effects of *In Utero* Influenza Exposure in the Post-1940 U.S. Population," *Journal of Political Economy*, 114: 672-712.

Andersson, Jens and Berit Carlstedt (2003), *Urval till Plikttjänst*, ILM Serie T:28. Karlstad, Sweden: Institutionen för ledarskap och management, Försvarshögskolan.

Behrman, Jere R. and Mark R. Rosenzweig (2004), "Returns to Birth Weight," *Review of Economics and Statistics*, 86: 586-601.

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: 409-39.

Case, Anne, Angela Fertig and Christina Paxson (2005), "The lasting impact of childhood health and circumstance," *Journal of Health Economics*, 24: 365-89.

Case, Anne and Christina Paxson (2008), "Height, Health, and Cognitive Function at Older Ages," *American Economic Review*, 98: 463-7.

Chatterji, Pinka, Margarita Alegria, Mingshan Lu and David Takeuchi (2007), ”Psychiatric Disorders and Labor Market Outcomes: Evidence from the National Latino and Asian American Study," *Health Economics*, 16: 1069-90.

Claussen, Bjorgulf, G Davey Smith and Dag Thelle (2003), "Impact of Childhood and Adulthood Socioeconomic Position on Cause Specific Mortality: The Oslo Mortality Study", *Journal of Epidemiology and Community Health*, 57: 40-5.

Cook, Derek G., David P. Strachan (1999), "Summary of Effects of Parental Smoking on the Respiratory Health of Children and Implications for Further Research", *Thorax*, 54: 357-66.

Cunha, Flavio, James J. Heckman and Lance Lochner (2006), “Interpreting the Evidence on Life-Cycle Skill Formation”. In Handbook of the Economics of Education, Vol. 1, ed. Erik Hanushek and Finis Welch. Amsterdam, NL: North Holland.

Currie, Janet and Rosemary Hyson (1999), “Is the Impact of Health Shocks Cushioned by Socio-Economic Status? The Case of Low Birth weight,” *American Economic Review*, 89: 245-50.

Currie, Janet and Mark Stabile (2007), “Mental Health in Childhood and Human Capital”, NBER Chapters, in *The Problems of Disadvantaged Youth: An Economic Perspective*, 115-48, National Bureau of Economic Research (NBER).

Currie, Janet, Mark Stabile, Phongsack Manivong and Leslie L. Roos (2010), “Child Health and Young Adult Outcomes,” *Journal of Human Resources*, 45: 517-48.

Ettner, Susan L., Richard G. Frank, Ronald C. Kessler (1997), “The Impact of Psychiatric Disorders on Labor Market Outcomes,” *Industrial and Labor Relations Review*, 51: 64-81.

Ettner, Susan L., Johanna Catherine Maclean and Michael T. French (2011), “Does Having a Dysfunctional Personality Hurt Your Career? Axis II Personality Disorders and Labor Market Outcomes,” *Industrial Relations: A Journal of Economy and Society*, 50: 149-73.

Famulari, Melissa (1992), “The Effects of a Disability on Labor Market Performance: The Case of Epilepsy,” *Southern Economic Journal*, 58: 1072-87.

Fanning, Desmond (1981), “Ill Health Retirement as an Indicator of Morbidity,” *Journal of the Society of Occupational Medicine*, 31: 103-11.

Ferrari, Michel D. (1998), “The Economic Burden of Migraine to Society,” *Pharmacoeconomics*, 13: 667-76.

Krahn, Murray D., Catherine Berka, Peter Langois and Allan Detsky (1996), "Direct and Indirect Costs of Asthma in Canada," *Canadian Medical Association Journal*, 154: 821-54.

Lichtenstein, Paul, Patrick F. Sullivan, Sven Cnattingius, Margaret Gatz, Sofie Johansson, Eva Carlström, Camilla Björk, Magnus Svartengren, Alicja Volk, Lars Klareskog, Ulf de Faire, Martin Schalling, Juni Palmgren and Nancy L. Pedersen (2006), "The Swedish Twin Registry in the Third Millennium: An Update," *Twin Research and Human Genetics*, 9: 875-82.

Lindqvist, Erik and Roine Vestman (2009), "The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from the Swedish Enlistment," IFN Working Paper 794, Research Institute of Industrial Economics (IFN).

Lundborg, Petter, Anton Nilsson and Dan-Olof Rooth (2011), "Does Early Life Health Predict Schooling Within Twin Pairs?", IZA Working Paper 5803, Institute for the Study of Labor (IZA).

Lundborg, Petter, Paul Nystedt and Dan-Olof Rooth (2009), "The Height Premium in Earnings: The Role of Physical Capacity and Cognitive and Non-Cognitive Skills," IZA Working Paper 4266, Institute for the Study of Labor (IZA).

Malone, Daniel C., Kenneth A. Lawson, David H. Smith, H. Michael Arrighi and Carmelina Battista (1997), "A Cost of Illness Study of Allergic Rhinitis in the United States," *Journal of Allergy and Clinical Immunology*, 1: 22-7.

Ng, Ying Chu, Philip Jacobs and Jeffrey A. Johnson (2001), "Productivity Losses Associated with Diabetes in the U.S.," *Diabetes Care*, 24: 257-61.

Royer, Heather (2009), "Separated at Girth: US Twin Estimates of the Effects of Birth Weight," *American Economic Journal: Applied Economics*, 1: 49-85.

Salm, Martin and Daniel Schunk (2011), "The relationship between childhood health, developmental gaps, and parental education: Evidence from administrative data," *Journal of the European Economic Association*, forthcoming.

Smith, James P. (2009), "The impact of childhood health on adult labor market outcomes," *Review of Economics and Statistics*, 91: 478-89.

Socialstyrelsen (2008), *Dödsorsaker 2008*. Stockholm: Socialstyrelsen.

Ward, Elizabeth, Ahmedin Jemal, Vilma Cokkinides, Gopal K. Singh, Cheryll Cardinez, Asma Ghafoor and Michael Thun (2004), "Cancer Disparities by Race/Ethnicity and Socioeconomic Status", *CA: A Cancer Journal for Clinicians*, 54: 78-93.

WHO (1967), *Manual of the international statistical classification of diseases, injuries, and causes of death*, eight revision. Geneva: World Health Organization (WHO).

WHO (2004), *The World Health Report 2004*. Geneva: World Health Organization (WHO).

Tables

Table 1: Descriptive statistics for the full sample.

NON-BINARY VARIABLES	MEAN (STD)	MISSING OBSERVATIONS
(Log) earnings	12.39 (0.83)	-
Global health	2.07 (3.71)	-
Years of schooling	11.95 (2.24)	589
Cognitive ability	0.00 (1.00)	25,948
Noncognitive ability	0.00 (1.00)	17,730
Parental characteristics:		
Father's (log) earnings	11.65 (0.45)	25,176
Father's years of schooling	10.10 (2.37)	295,754
Mother's years of schooling	9.33 (2.44)	181,139
BINARY VARIABLES	MEAN	INCIDENCE
Any diagnose	0.507	359,691
Tumor	0.002	1,126
Endocrine, nutritional and metabolic	0.016	11,038
Mental	0.064	45,305
Nervous system	0.015	10,578
Sensory organs	0.152	107,767
Circulatory	0.027	19,367
Respiratory	0.121	86,074
Digestive	0.030	21,493
Genito-urinary	0.014	10,032
Skin	0.054	38,317
Musculoskeletal and connective tissue	0.136	96,530
Congenital anomalies	0.017	11,988
Injuries and poisonings	0.049	34,612
Diabetes	0.002	1,479
Neurosis	0.028	19,972
Personality disorder	0.008	5,712
Alcoholism and drug dependence	0.005	3,273
Epilepsy	0.004	2,497
Migraine	0.009	6,605
Asthma	0.027	19,404
Hay fever	0.076	54,084
Vertebrogenic pain syndrome	0.054	38,240
Number of observations	*710,018*	

Table 2: Descriptive statistics for the siblings subsample.

NON-BINARY VARIABLES	MEAN (STD)	MISSING OBSERVATIONS
(Log) earnings	12.39 (0.81)	-
Global health	2.03 (3.69)	-
Years of schooling	11.82 (2.23)	255
Cognitive ability	0.00 (0.98)	9,695
Noncognitive ability	0.00 (0.99)	6,320
Parental characteristics:		
Father's (log) earnings	11.66 (0.45)	7,005
Father's years of schooling	9.93 (2.44)	107,987
Mother's years of schooling	9.73 (2.52)	55,443
BINARY VARIABLES	MEAN	INCIDENCE
Any diagnose	0.497	135,882
Tumors	0.002	412
Endocrine, nutritional and metabolic	0.014	3,834
Mental	0.063	17,335
Nervous system	0.014	3,881
Sensory organs	0.156	42,541
Circulatory	0.026	7,162
Respiratory	0.107	29,362
Digestive	0.030	8,310
Genito-urinary	0.013	3,594
Skin	0.052	14,237
Musculoskeletal and connective tissue	0.135	36,793
Congenital anomalies	0.016	4,385
Injuries and poisonings	0.047	12,789
Diabetes	0.002	526
Neurosis	0.027	7,389
Personality disorder	0.008	2,210
Alcoholism and drug dependence	0.005	1,260
Epilepsy	0.004	984
Migraine	0.009	2,395
Asthma	0.024	6,642
Hay fever	0.066	17,958
Vertebrogenic pain syndrome	0.056	15,285
Number of observations	*273,296*	

Table 3. The (bad) health penalty for *Global health*. Men born 1950-1970. Logarithm of annual earnings in 2003. Full sample.

Variable	A	B	C	D	E	F	G	H
Global health(*10):	-0.250***	-0.229***	-0.174***	-0.095***	-0.087***	-0.190***	-0.153***	-0.075***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Parental characteristics:								
Father's (log) earnings	-	0.232***	0.179***	0.196***	0.160***	0.148***	-	0.077***
		(0.003)	(0.003)	(0.003)	(0.003)	(0.003)		(0.002)
Father's years of schooling	-	0.011***	0.003***	0.007***	0.002***	0.000	-	-0.004***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.001)		(0.001)
Mother's years of schooling	-	0.012***	0.004***	0.008***	0.002***	0.000	-	-0.003***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.001)		(0.001)
Cognitive skill:	-	-	0.141***	-	0.113***	-	-	0.027***
			(0.001)		(0.001)			(0.001)
Noncognitive skill:	-	-	-	0.132***	0.099***	-	-	0.060***
				(0.001)	(0.001)			(0.001)
Schooling:	-	-	-	-	-	0.173***	-	0.054***
						(0.001)		(0.001)
Occupation fixed effects	No	No	No	No	No	No	Yes	Yes
R2	0.01	0.04	0.06	0.06	0.07	0.07	0.21	0.22
Number of observations	710,018							

Notes: Columns A through G report estimates from the (1) regression model: Log $Earnings_i = a + b*GHt_i + c*X_i + d*F_i + e_i$ using the total population data. Model A only has GH and age fixed effects and is estimated using OLS. Model B adds the parental variables. Model C adds cognitive skill and Model D noncognitive skill, respectively, to Model B, while Model E adds both variables to Model B. Model F adds years of schooling while Model G adds occupation fixed effects to Model A. All regressions were run in Stata 11 using robust standard errors. Fixed effects for birth cohort were included.

Table 4. The bad health penalty for *Global health*. Men born 1950-1970. Logarithm of annual earnings in 2003. Siblings.

Variable	A	B	C	D	E	F	G	H
Global health(*10):	-0.252***	-0.175***	-0.137***	-0.097***	-0.086***	-0.155***	-0.131***	-0.077***
	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)
Cognitive skill:	-	-	0.131***	-	0.110***	-	-	0.040***
			(0.003)		(0.003)			(0.003)
Noncognitive skill:	-	-	-	0.100***	0.075***	-	-	0.050***
				(0.003)	(0.003)			(0.003)
Schooling	-	-	-	-	-	0.148***	-	0.052***
						(0.003)		(0.001)
Occupation fixed effects	No	No	No	No	No	No	Yes	Yes
Sibling fixed effects	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.01	0.02	0.05	0.04	0.07	0.07	0.20	0.21
No of observations	273,296							

Notes: Columns B through H report estimates from the (2) regression model: Log $Earnings_{ij} = a + b*GH_{ij} + c*X_{ij} + f_j + e_{ij}$ using the sibling subsample. Model A contains GH and age fixed effects for the sibling sample and is estimated using OLS. Model B adds sibling fixed effects and is estimated using the xtreg command in Stata 11. Model C adds cognitive skill and Model D noncognitive skill, respectively, to Model B, while Model E adds both variables to Model B. Model F adds years of schooling to Model B, while Model G adds an occupation fixed effect. All regressions were run using robust standard errors. Fixed effects for birth cohort were included.

Table 5: The (bad) health penalty for *Global health*. Men born 1950-1970. Logarithm of annual earnings in 2003. Twins.

Variable	A	B	C	D	E	F
Global health(*10):	-0.228***	-0.119***	-0.189***	-0.157***	-0.218***	-0.135*
	(0.030)	(0.042)	(0.038)	(0.053)	(0.059)	(0.077)
Type of twin	MZ+DZ+UK	MZ+DZ+UK	DZ	DZ	MZ	MZ
Twin fixed effects	No	Yes	No	Yes	No	Yes
R2	0.01	0.01	0.01	0.01	0.05	0.02
No of observations	6,867		3,185		2,316	

Notes: The above models contain GH and no controls.

Table 6. The (bad) health penalty for *Broad categories of health conditions*. Men born 1950-1970. Logarithm of annual earnings in 2003. Full sample.

VARIABLES	A	B	C	D	E	F	G	H
Diagnose=1:	-0.115***	-0.104***	-0.080***	-0.039***	-0.037***	-0.084***	-0.068***	-0.031***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
By type:								
Tumors	-0.040*	-0.037	-0.021	0.008	0.006	-0.036	-0.023	0.002
	(0.025)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.022)	(0.022)
Endocrine etc.	-0.171***	-0.147***	-0.107***	-0.077***	-0.067***	-0.114***	-0.098***	-0.053***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)
Mental	-0.314***	-0.288***	-0.212***	-0.106***	-0.093***	-0.234***	-0.185***	-0.079***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Nervous system	-0.118***	-0.113***	-0.091***	-0.059***	-0.057***	-0.100***	-0.081***	-0.051***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)
Sensory organs	-0.055***	-0.045***	-0.031***	-0.026***	-0.019***	-0.037***	-0.027***	-0.015***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
Circulatory	0.001	-0.002	-0.003	-0.001	-0.002	-0.008	-0.007	-0.009
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)
Respiratory	0.011***	0.001	-0.010***	0.016***	0.003	-0.008***	-0.010***	-0.008***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Digestive	-0.061***	-0.051***	-0.043***	-0.018***	-0.020***	-0.041***	-0.037***	-0.017***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Geniro-urinary	-0.028***	-0.025***	-0.024***	-0.015*	-0.017**	-0.029***	-0.020***	-0.016**
	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Skin	-0.008*	-0.011**	-0.014***	0.002	-0.004	-0.013***	-0.006	-0.002
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Musculoskeletal etc.	-0.069***	-0.057***	-0.046***	-0.034***	-0.030***	-0.040***	-0.035***	-0.020***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Congentital anomalies	-0.048***	-0.046***	-0.033***	-0.012	-0.012	-0.044***	-0.034***	-0.016**
	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)
Injuries and poisonings	-0.044***	-0.035***	-0.025***	-0.026***	-0.020***	-0.018***	-0.021***	-0.012***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)
Parental characteristics								

Father's (log) earnings		0.229***	0.177***	0.195***	0.162***	0.146***		0.076***
		(0.003)	(0.003)	(0.003)	(0.003)	(0.003)		(0.002)
Father's years of schooling		0.011***	0.003***	0.007***	0.002***	-0.000		-0.004***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)		(0.000)
Mother's years of schooling		0.012***	0.004***	0.008***	0.002***	-0.000		-0.003***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)		(0.000)
Cognitive skill			0.139***		0.112***			0.027***
			(0.001)		(0.001)			(0.001)
Noncognitive skill				0.129***	0.096***			0.058***
				(0.001)	(0.001)			(0.001)
Schooling						0.172***		0.054***
						(0.001)		(0.001)
Occupation fixed effects	No	No	No	No	No	No	Yes	Yes
R2	0.016	0.038	0.062	0.058	0.072	0.074	0.207	0.219
No of observations				710,018				

Notes: The parameter estimates for the X variables refer to the model when diagnoses are entered by type. All regressions were run in Stata 11 using robust standard errors. Fixed effects for birth cohort were included.

Table 7. The (bad) health penalty for *Broad categories of health conditions*. Men born 1950-1970. Logarithm of annual earnings in 2003. Siblings.

VARIABLES	A	B	C	D	E	F	G	H
Diagnose=1:	-0.115***	-0.104***	-0.080***	-0.039***	-0.037***	-0.084***	-0.068***	-0.031***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
By type:								
Tumors	0.010	-0.028	-0.015	0.003	0.004	-0.018	-0.003	0.019
	(0.035)	(0.051)	(0.050)	(0.050)	(0.050)	(0.050)	(0.047)	(0.047)
Endocrine etc.	-0.180***	-0.104***	-0.080***	-0.060***	-0.054***	-0.091***	-0.077***	-0.043***
	(0.015)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.016)	(0.016)
Mental	-0.314***	-0.202***	-0.153***	-0.094***	-0.080***	-0.177***	-0.147***	-0.074***
	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)
Nervous system	-0.112***	-0.085***	-0.069***	-0.047***	-0.044***	-0.077***	-0.068***	-0.043***
	(0.014)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.016)	(0.016)
Sensory organs	-0.057***	-0.034***	-0.026***	-0.024***	-0.020***	-0.031***	-0.022***	-0.016***
	(0.004)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)
Circulatory	0.005	0.013	0.010	0.013	0.011	0.013	0.010	0.009
	(0.010)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.011)
Respiratory	0.008	-0.016**	-0.019***	-0.005	-0.010	-0.018***	-0.020***	-0.015**
	(0.005)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)
Digestive	-0.045***	-0.013	-0.011	0.005	0.002	-0.009	-0.006	0.004
	(0.009)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.011)	(0.011)
Genito-urinary	-0.038***	-0.027	-0.025	-0.018	-0.018	-0.026	-0.023	-0.018
	(0.014)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.016)	(0.016)
Skin	-0.009	-0.015	-0.012	-0.008	-0.007	-0.012	-0.005	-0.001
	(0.007)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)
Musculoskeletal etc.	-0.067***	-0.024***	-0.018***	-0.011*	-0.009	-0.016***	-0.014***	-0.005
	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Congenital anomalies	-0.024**	-0.046***	-0.033**	-0.027*	-0.021	-0.043***	-0.044***	-0.028*
	(0.012)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.015)	(0.015)
Injuries and poisonings	-0.045***	-0.040***	-0.034***	-0.031***	-0.028***	-0.033***	-0.030***	-0.023***
	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Cognitive skill			0.131***		0.110***			0.040***
			(0.003)		(0.003)			(0.003)

Noncognitive skill				0.100***	0.075***			0.050***
				(0.003)	(0.003)			(0.002)
Schooling						0.148***		0.052***
						(0.003)		(0.003)
Occupation fixed effects	No	No	No	No	No	No	Yes	Yes
Sibling fixed effects	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.017	0.008	0.023	0.019	0.029	0.029	0.148	0.156
No of observations				273,296				

Notes: The parameter estimates for the X variables refer to the model when diagnoses are entered by type. All regressions were run in Stata 11 using robust standard errors. Fixed effects for birth cohort were included.

Table 8. The (bad) health penalty for *Specific types of health conditions*. Men born 1950-1970. Logarithm of annual earnings in 2003. Full sample.

Variable	A	B	C	D	E	F	G	H
Diabetes	-0.178***	-0.182***	-0.107***	-0.010	-0.020	-0.189***	-0.167***	-0.061**
	(0.027)	(0.027)	(0.028)	(0.028)	(0.028)	(0.027)	(0.025)	(0.026)
Neurosis	-0.288***	-0.279***	-0.223***	-0.089***	-0.095***	-0.239***	-0.186***	-0.084***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)
Personality disorder	-0.331***	-0.302***	-0.233***	-0.117***	-0.111***	-0.255***	-0.207***	-0.102***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.014)	(0.014)
Alcoholism and drugs	-0.409***	-0.381***	-0.317***	-0.180***	-0.183***	-0.310***	-0.234***	-0.120***
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.018)	(0.018)
Epilepsy	-0.152***	-0.150***	-0.095***	-0.087***	-0.063***	-0.132***	-0.099***	-0.063***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.018)	(0.018)
Migraine	-0.034***	-0.033***	-0.040***	-0.010	-0.021**	-0.032***	-0.031***	-0.021***
	(0.011)	(0.011)	(0.010)	(0.011)	(0.010)	(0.011)	(0.010)	(0.010)
Asthma	-0.009	-0.007	-0.010*	0.012*	0.003	-0.009	-0.012**	-0.002
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Hay fever	0.091***	0.065***	0.036***	0.049***	0.031***	0.039***	0.032***	0.013***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)
Vertebrogenic pain	-0.082***	-0.063***	-0.050***	-0.040***	-0.034***	-0.040***	-0.040***	-0.021***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)
Parental characteristics:								
Father's (log) earnings	-	0.231***	0.178***	0.195***	0.162***	0.147***	-	0.076***
		(0.003)	(0.003)	(0.003)	(0.003)	(0.003)		(0.002)
Father's years of schooling	-	0.011***	0.003***	0.007***	0.002***	0.000	-	-0.004***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.001)		(0.001)
Mother's years of schooling	-	0.012***	0.004***	0.008***	0.002***	0.000	-	-0.003***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.001)		(0.001)
Cognitive skill:	-	-	0.141***	-	0.113***	-	-	0.027***
			(0.001)		(0.001)			(0.001)
Noncognitive skill:	-	-	-	0.132***	0.098***	-	-	0.060***
				(0.001)	(0.001)			(0.001)
Schooling	-	-	-	-	-	0.173***	-	0.053***
						(0.001)		(0.001)

Occupation fixed effects	No	No	No	No	No	No	Yes	Yes
R2	0.01	0.03	0.06	0.06	0.07	0.07	0.21	0.22
No of observations				710,018				

Notes: All regressions were run in Stata 11 using robust standard errors. Fixed effects for birth cohort were included.

Table 9. The (bad) health penalty for *Specific types of health conditions*. Men born 1950-1970. Logarithm of annual earnings in 2003. Siblings.

Variable	A	B	C	D	E	F	G	H
Diabetes	-0.165***	-0.161***	-0.099**	-0.043	-0.038**	-0.158***	-0.153***	-0.056
	(0.046)	(0.046)	(0.046)	(0.047)	(0.047)	(0.045)	(0.042)	(0.043)
Neurosis	-0.298***	-0.206***	-0.170***	-0.095***	-0.093***	-0.186***	-0.157***	-0.086***
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.012)	(0.012)
Personality disorder	-0.316***	-0.188***	-0.143***	-0.079***	-0.070***	-0.164***	-0.131***	-0.060***
	(0.023)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.021)	(0.021)
Alcoholism and drugs	-0.416***	-0.265***	-0.227***	-0.143***	-0.143***	-0.240***	-0.185***	-0.109***
	(0.032)	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)	(0.028)	(0.028)
Epilepsy	-0.140***	-0.137***	-0.091***	-0.090***	-0.066***	-0.130***	-0.112***	-0.074***
	(0.030)	(0.034)	(0.033)	(0.033)	(0.033)	(0.033)	(0.031)	(0.031)
Migraine	-0.033*	-0.002	-0.010	0.014	0.004	0.000	-0.008	-0.003
	(0.017)	(0.022)	(0.021)	(0.022)	(0.021)	(0.021)	(0.020)	(0.020)
Asthma	0.002	0.000	-0.002	0.014	0.008	-0.002	-0.014	-0.007
	(0.011)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.012)
Hay fever	0.087***	0.017**	0.007	0.010	0.004	0.010	0.005	-0.003
	(0.007)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)
Vertebrogenic pain	-0.067***	-0.004	0.000	0.006	0.008	0.004	-0.001	0.008
	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)
Cognitive skill:	-	-	0.132***	-	0.110***	-	-	0.041***
			(0.003)		(0.003)			(0.003)
Noncognitive skill:	-	-	-	0.102***	0.076***	-	-	0.051***
				(0.002)	(0.003)			(0.002)
Schooling	-	-	-	-	-	0.149***	-	0.052***
						(0.003)		(0.000)
Occupation fixed effects	No	No	No	No	No	No	Yes	Yes
Sibling fixed effects	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.01	0.01	0.05	0.04	0.07	0.07	0.20	0.21
No of observations	273,296							

Notes: All regressions were run in Stata 11 using robust standard errors. Fixed effects for birth cohort were included.