

IZA DP No. 5804

**Early Life Health and Adult Earnings:
Evidence from a Large Sample of Siblings and Twins**

Petter Lundborg
Anton Nilsson
Dan-Olof Rooth

June 2011

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

Petter Lundborg

*Lund University, VU University Amsterdam,
Tinbergen Institute, Netspar, HEP and IZA*

Anton Nilsson

Lund University

Dan-Olof Rooth

*Linnaeus University,
Lund University, CReAM and IZA*

Discussion Paper No. 5804
June 2011

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

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

We study the relationship between early life health 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 diagnoses have long-run effects on future earnings with the largest effects resulting from mental conditions. 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

Corresponding author:

Dan-Olof Rooth
Linnaeus School of Business and Economics
Linnaeus University
Kalmar Nyckel
SE-39182 Kalmar
Sweden
E-mail: Dan-Olof.Rooth@lnu.se

1. Introduction

An adult's success on the labor market is to a large extent already determined during his childhood. This is suggested by a growing literature that links childhood health to various 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 disproportionately to be found in families of low socioeconomic status, this implies that health during childhood 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 childhood health on adult outcomes is clearly needed, however.

In this paper, we study the relationship between early life health 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 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 physicians. 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.

Fourth, our data allow 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. We are unaware of any previous study estimating the effect of childhood health after birth 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 or a condition of the nervous system at age 18 gives rise to an earnings penalty of about 20 percent when sibling fixed effects are included. For both these types of conditions, the estimates drop to about 10 percent when controls for cognitive and noncognitive ability are added. This may be viewed as a lower bound, however, as both cognitive and noncognitive ability may be partly determined by early life health status.

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 health. Using data from the 1958 British cohort study, Case et al. (2005) showed that adults who had

suffered from chronic conditions during childhood had lower educational attainment, wages, and employment probabilities compared to other adults.

Smith (2009) used a retrospective measure of self-assessed health at 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. 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 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 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 of the variation in cognitive and verbal test scores. These results are important, because they point to one mechanism whereby early health may affect later socioeconomic achievement.

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.

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 individuals with severe physical or mental handicaps, prisoners, and individuals living abroad. Individuals usually took the test at the age of 18 or 19. The military enlistment records include up to six medical conditions for every individual, using the ICD-8 classification. 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.

It is an advantage of our data that health is determined based on obligatory assessments conducted by a physician. While these assessments by necessity partly rest on self-reports, 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, are less of an issue. The main disadvantage is that conditions that are too mild to affect the individual's suitability for military service may not have been recorded.

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 jäger and 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 and shoe repairing). Individuals assigned a “Y” or “Z” (in total 8 %) are not allowed to undergo education within the military.² “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 14 broad classes of health conditions, which generally follow the main classes provided by the ICD-8 categorization (WHO, 1967). These 14 groups of health conditions cover more than 98 percent of all health problems reported in our “full sample” (as defined below). In creating these classes of health diagnoses, we deviate from the ICD-8 classification only in two respects. First, instead of “circulatory conditions” we have included one category of heart conditions³ and one category of “other circulatory conditions”.⁴ Second, 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 system⁵ and one category of conditions of the sensory organs⁶. This leaves us with the following classes of health conditions: tumors, endocrine

² However, individuals assigned a “Y” may be required to undergo non-military education.

³ ICD-8: 393-429.

⁴ ICD-8: 430-458. This includes cerebrovascular diseases, diseases of arteries, arterioles and capillaries, diseases of veins and lymphatics and other diseases of the circulatory system.

⁵ ICD-8: 320-358.

⁶ ICD-8: 360-389.

conditions, mental conditions, conditions of the nervous system, conditions of the sensory organs, heart conditions, other circulatory conditions, respiratory conditions, digestive conditions, genito-urinary conditions, skin conditions, musculoskeletal conditions, congenital anomalies and injuries and poisonings.

In addition to these broad classes of health problems, we analyze a number of specific conditions. 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 musculoskeletal condition). Previous literature has documented relations between all these diagnoses and adverse labor market outcomes, however in most cases only simultaneous relationships using sources like sickness absence records or standard self-reports.⁷ 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 is 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.⁸ 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

⁷ 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).

⁸ We treat zeros as missing since individuals should get at least a few points by pure chance.

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 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 SSK (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 through 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, health conditions were reported by the National Service Administration using the ICD-9 instead of the ICD-8 classification. Since ICD-8 and ICD-9 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 have not undergone the military enlistment procedure or for other reasons lack data on (global) health status (18.6 %), our sample includes 774,529 individuals. Second, individuals not reported to have positive annual earnings in 2003 are excluded, 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 Registry⁹ (“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 control for genetic factors.

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 siblings 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. Musculoskeletal conditions and conditions of the sensory organs are the largest groups, followed by respiratory conditions.

3.3. Empirical method

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

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

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

where i is index for individual. H is either a scalar representing global health, a vector of the different classes of health conditions (including an indicator for “other classes”), 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 poor health *partially* determines other individual characteristics and that no other confounding mechanisms are at play, estimates in models without controls for individual characteristics can be seen as upper bounds whereas estimates including such controls can be seen as lower bounds of the effects of health 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:

$$2) \quad \text{Log Earnings}_{ij} = a + b * H_{ij} + c * X_{ij} + d * f_i + e_{ij}.$$

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.

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 health than required for military service, implies a 25 percent decrease in earnings. Given 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 a possible lower bound of 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. This means that there is 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 impact 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 (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.

4.2 Global health – the siblings 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.

Including cognitive or noncognitive ability again substantially lowers the health penalty (Models C-E). In particular, after including noncognitive ability, estimates become very close 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-income relationship is not that individuals in poor health work in

other occupations than healthy people, but instead that individuals in poor health acquire less well-paid jobs within 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 such as cognitive and noncognitive ability (not reported). The inclusion of these variables was again found to reduce the health penalty in a similar magnitude as what we found for the full sample and for siblings. However, the health penalty became insignificant at the 10 percent level in the specifications including both cognitive and noncognitive ability. Including cognitive but not noncognitive ability kept the health penalty significant among DZ twins, but not among MZ ones. These results imply that we cannot fully reject the possibility that health has no direct effect on earnings. However, as long as cognitive and noncognitive ability is affected by health status, an indirect effect is of course still present.

4.4 Classes of health conditions

Regression results for specific classes 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 representative of the larger sample.

As can be seen, nearly all types of conditions are strongly statistically significantly associated with lower earnings, irrespective of the specification. Tumors and circulatory conditions are the exceptions, as these often 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 167 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 effects of tumors and circulatory conditions (at least heart conditions) on adult earnings.

The inclusion of sibling fixed effects in general reduces the health penalties by between 10 and 50 percent (Model B in Table 7 compared to Model A). Among the significant diagnoses, the exception to this pattern is respiratory conditions for which the estimate instead increases by about 10 percent, indicating that individuals with respiratory conditions on average come from advantaged family backgrounds. A similar conclusion can be drawn from Table 6, where the coefficient on respiratory conditions is the only to increase when controls for parental characteristics are added (Model B).

In addition to specific classes 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 estimate is reduced to a negative of 7 percent, and controlling for cognitive and noncognitive ability further reduces it to a negative of 3 percent. Again, the health penalty is to the largest

part driven by individuals in poor health acquiring less well-paid jobs within the same occupations compared to healthy people. This pattern also emerges for all classes of health conditions, and for all specific health diagnoses with significant effects (see below).

Some classes of health conditions have very large effects. In particular, endocrine conditions, mental conditions and conditions of the nervous system have coefficients similar to or higher than the ones obtained for schooling, cognitive ability and noncognitive ability. We find that almost 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 24 percent less, and according to Model G which also includes occupation fixed effects, these individuals earn on average 17 percent less. Only in some of the models including noncognitive ability this pattern is broken and conditions of the nervous system have more severe effects. These models may be less meaningful, however, as our measure of noncognitive ability is likely to be quite strongly endogenous to mental illness.

Conditions of the nervous system stands out as the second most severe type of condition with a coefficient of -0.18 when sibling fixed effects are included (Model B in Table 7), -0.15 when occupation fixed effects are added (Model G), and -0.11 when cognitive ability, noncognitive ability and schooling are also included (Model H). As number three, endocrine conditions turn up with coefficients between -0.14 when only sibling effects are included (Model B) and -0.045 when full controls are added (Model H).

Other classes of conditions have generally much smaller effects. For example, conditions of the sensory organs only have between one-fourth and one-fifth of the effect of mental conditions, depending on the specification. Injuries and congenital anomalies, which were for example examined by Currie et al. (2010), tend to have effects between one-third and one-half of the effect of mental conditions. Still, these conditions imply an eight percent

earnings reduction in the model including sibling fixed effects (Model B in Table 7) and a four percent earnings reduction in the model including full controls (Model H in Table 7).

4.5 Specific health diagnoses

Table 8-9 show our results for the specific 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 (ability, years of schooling and occupation fixed effects) are added. This effect is still twice as large as the effects of cognitive skill, noncognitive skill, or schooling (Model H).¹⁰

Besides alcoholism and drug dependence, personality disorder and neurosis tend to have the largest coefficients. The negative effects of both these 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 larger percentage than the coefficient of neurosis and the coefficient on alcoholism and drug dependence when controls are added, indicating that either confounding characteristics related

¹⁰ Given that alcoholism and drug dependence are conditions that are likely to develop quite late in adolescence when IQ and other personal characteristics have already stabilized, the argument for including controls such as cognitive and noncognitive ability is rather strong in this case.

to personality disorders or indirect paths between personality disorders and earnings are comparatively important.

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 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 early health 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 classes of health conditions, as well as specific diagnoses. This analysis shows us that while most classes of diseases seem to have important long-run effects, the strongest ones are detected for mental conditions, with rather similar effects of the particular 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.

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, 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, 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.

WHO (1967), *Manual of the international statistical classification of diseases, injuries, and causes of death*, eight revision. 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
<i>Parental characteristics:</i>		
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
Tumors	0.001	424
Endocrine	0.007	4,692
Mental	0.026	18,512
Nervous system	0.007	5,031
Sensory organs	0.094	67,049
Heart	0.013	9,098
Other circulatory	0.004	2,896
Respiratory	0.080	56,700
Digestive	0.019	13,768
Genito-urinary	0.010	7,110
Skin	0.041	29,368
Musculoskeletal	0.123	87,476
Congenital anomalies	0.016	11,401
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
<i>Number of observations</i>	<i>710,018</i>	

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
<i>Parental characteristics:</i>		
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	0.001	167
Endocrine	0.006	1,659
Mental	0.026	7,192
Nervous system	0.007	1,864
Sensory organs	0.099	27,041
Heart	0.013	3,515
Other circulatory	0.004	1,080
Respiratory	0.071	19,291
Digestive	0.020	5,363
Genito-urinary	0.009	2,592
Skin	0.040	11,019
Musculoskeletal	0.123	33,529
Congenital anomalies	0.015	4,176
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
<i>Number of observations</i>	<i>273,296</i>	

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*** (0.003)	-0.229*** (0.003)	-0.174*** (0.003)	-0.095*** (0.003)	-0.087*** (0.003)	-0.190*** (0.003)	-0.153*** (0.003)	-0.075*** (0.003)
<i>Parental characteristics:</i>								
Father's (log) earnings	-	0.232*** (0.003)	0.179*** (0.003)	0.196*** (0.003)	0.160*** (0.003)	0.148*** (0.003)	-	0.077*** (0.002)
Father's years of schooling	-	0.011*** (0.000)	0.003*** (0.000)	0.007*** (0.000)	0.002*** (0.000)	0.000 (0.001)	-	-0.004*** (0.001)
Mother's years of schooling	-	0.012*** (0.000)	0.004*** (0.000)	0.008*** (0.000)	0.002*** (0.000)	0.000 (0.001)	-	-0.003*** (0.001)
Cognitive skill:	-	-	0.141*** (0.001)	-	0.113*** (0.001)	-	-	0.027*** (0.001)
Noncognitive skill:	-	-	-	0.132*** (0.001)	0.099*** (0.001)	-	-	0.060*** (0.001)
Schooling:	-	-	-	-	-	0.173*** (0.001)	-	0.054*** (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: $\text{Log Earnings}_i = a + b \cdot \text{GH}_i + c \cdot X_i + d \cdot F_i + e_i$ 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*** (0.005)	-0.175*** (0.006)	-0.137*** (0.006)	-0.097*** (0.006)	-0.086*** (0.006)	-0.155*** (0.006)	-0.131*** (0.005)	-0.077*** (0.005)
Cognitive skill:	-	-	0.131*** (0.003)	-	0.110*** (0.003)	-	-	0.040*** (0.003)
Noncognitive skill:	-	-	-	0.100*** (0.003)	0.075*** (0.003)	-	-	0.050*** (0.003)
Schooling	-	-	-	-	-	0.148*** (0.003)	-	0.052*** (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: $\text{Log Earnings}_{ij} = a + b \cdot \text{GH}_{ij} + c \cdot X_{ij} + f_j + e_{ij}$ 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*** (0.030)	-0.119*** (0.042)	-0.189*** (0.038)	-0.157*** (0.053)	-0.218*** (0.059)	-0.135* (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 *Health diagnoses*. Men born 1950-1970. Logarithm of annual earnings in 2003. Full sample.

Variable	A	B	C	D	E	F	G	H
Diagnose=1:	-0.115*** (0.002)	-0.104*** (0.002)	-0.080*** (0.002)	-0.039*** (0.002)	-0.037*** (0.002)	-0.084*** (0.002)	-0.068*** (0.003)	-0.031*** (0.002)
<i>By type:</i>								
1 Tumors	-0.077** (0.037)	-0.059* (0.036)	-0.032 (0.036)	0.005 (0.036)	0.006 (0.036)	-0.060* (0.036)	-0.043 (0.033)	-0.004 (0.033)
2 Endocrine	-0.230*** (0.014)	-0.203*** (0.014)	-0.146*** (0.014)	-0.086*** (0.014)	-0.079*** (0.014)	-0.165*** (0.014)	-0.140*** (0.014)	-0.066*** (0.013)
3 Mental	-0.382*** (0.008)	-0.347*** (0.008)	-0.256*** (0.008)	-0.133*** (0.008)	-0.118*** (0.008)	-0.276*** (0.008)	-0.218*** (0.007)	-0.090*** (0.007)
4 Nervous system	-0.217*** (0.014)	-0.208*** (0.013)	-0.164*** (0.013)	-0.112*** (0.013)	-0.107*** (0.013)	-0.178*** (0.013)	-0.149*** (0.012)	-0.093*** (0.012)
5 Sensory organs	-0.088*** (0.003)	-0.074*** (0.003)	-0.050*** (0.003)	-0.033*** (0.003)	-0.026*** (0.003)	-0.059*** (0.003)	-0.044*** (0.003)	-0.019*** (0.003)
6 Heart	-0.020** (0.008)	-0.019** (0.008)	-0.009 (0.008)	0.002 (0.008)	0.004 (0.008)	-0.020*** (0.008)	-0.018** (0.007)	-0.006** (0.007)
7 Other circulatory	-0.079*** (0.017)	-0.082*** (0.017)	-0.074*** (0.017)	-0.054*** (0.017)	-0.055*** (0.017)	-0.077*** (0.017)	-0.062*** (0.015)	-0.052*** (0.015)
8 Respiratory	-0.032*** (0.004)	-0.038*** (0.004)	-0.039*** (0.004)	0.003 (0.004)	-0.008** (0.004)	-0.040*** (0.004)	-0.035*** (0.003)	-0.018*** (0.003)
9 Digestive	-0.131*** (0.008)	-0.114*** (0.008)	-0.090*** (0.007)	-0.038*** (0.008)	-0.039*** (0.007)	-0.093*** (0.007)	-0.077*** (0.007)	-0.032*** (0.007)
10 Genito-urinary	-0.091*** (0.010)	-0.084*** (0.010)	-0.073*** (0.010)	-0.043*** (0.010)	-0.044*** (0.010)	-0.080*** (0.010)	-0.060*** (0.009)	-0.038*** (0.009)
11 Skin	-0.060*** (0.005)	-0.060*** (0.005)	-0.053*** (0.005)	-0.013*** (0.005)	-0.019*** (0.005)	-0.054*** (0.005)	-0.040*** (0.005)	-0.018*** (0.004)
12 Musculoskeletal	-0.141*** (0.003)	-0.123*** (0.003)	-0.096*** (0.003)	-0.056*** (0.003)	-0.052*** (0.003)	-0.095*** (0.003)	-0.079*** (0.002)	-0.039*** (0.003)
13 Congenital anomalies	-0.118*** (0.008)	-0.109*** (0.008)	-0.082*** (0.008)	-0.036*** (0.008)	-0.035*** (0.008)	-0.095*** (0.008)	-0.074*** (0.007)	-0.034*** (0.007)
14 Injuries and poisonings	-0.125*** (0.005)	-0.108*** (0.005)	-0.081*** (0.005)	-0.054*** (0.005)	-0.047*** (0.005)	-0.077*** (0.005)	-0.070*** (0.004)	-0.034*** (0.004)
<i>Parental characteristics:</i>								
Father's (log) earnings	-	0.233*** (0.003)	0.178*** (0.003)	0.195*** (0.003)	0.161*** (0.003)	0.148*** (0.003)	-	0.076*** (0.002)
Father's years of schooling	-	0.011*** (0.000)	0.003*** (0.000)	0.007*** (0.000)	0.002*** (0.000)	0.000 (0.001)	-	-0.004*** (0.001)

Mother's years of schooling	-	0.012*** (0.000)	0.003*** (0.000)	0.008*** (0.000)	0.002*** (0.000)	0.000 (0.001)	-	-0.003*** (0.001)
Cognitive skill:	-	-	0.144*** (0.001)	-	0.112*** (0.001)	-	-	0.027*** (0.001)
Noncognitive skill:	-	-	-	0.136*** (0.001)	0.102*** (0.001)	-	-	0.064*** (0.001)
Schooling	-	-	-	-	-	0.174*** (0.001)	-	0.053*** (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	0.20	0.22
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 *Health diagnoses*. Men born 1950-1970. Logarithm of annual earnings in 2003. Siblings.

Variable	A	B	C	D	E	F	G	H
Diagnose=1:	-0.118*** (0.003)	-0.071*** (0.004)	-0.055*** (0.004)	-0.034*** (0.004)	-0.031*** (0.004)	-0.061*** (0.004)	-0.051*** (0.004)	-0.026*** (0.004)
<i>By type:</i>								
1 Tumors	-0.035 (0.048)	-0.061 (0.080)	-0.031 (0.079)	-0.016 (0.079)	-0.005 (0.079)	-0.060 (0.079)	-0.020 (0.074)	0.009 (0.073)
2 Endocrine	-0.229*** (0.022)	-0.136*** (0.026)	-0.100*** (0.026)	-0.063** (0.026)	-0.056** (0.026)	-0.115*** (0.026)	-0.101*** (0.024)	-0.045*** (0.024)
3 Mental	-0.382*** (0.012)	-0.240*** (0.013)	-0.183*** (0.013)	-0.120*** (0.013)	-0.104*** (0.013)	-0.208*** (0.013)	-0.171*** (0.012)	-0.085*** (0.012)
4 Nervous system	-0.216*** (0.022)	-0.184*** (0.024)	-0.155*** (0.024)	-0.125*** (0.024)	-0.118*** (0.024)	-0.170*** (0.024)	-0.154*** (0.022)	-0.112*** (0.022)
5 Sensory organs	-0.092*** (0.005)	-0.050*** (0.007)	-0.037*** (0.007)	-0.029*** (0.007)	-0.024*** (0.007)	-0.045*** (0.007)	-0.036*** (0.007)	-0.021*** (0.007)
6 Heart	-0.015 (0.012)	-0.006 (0.018)	-0.002 (0.018)	0.007 (0.018)	0.007 (0.017)	-0.004 (0.017)	0.004 (0.016)	0.011 (0.016)
7 Other circulatory	-0.060** (0.026)	-0.033 (0.032)	-0.028 (0.031)	-0.020 (0.031)	-0.019 (0.031)	-0.027 (0.031)	-0.038 (0.029)	-0.030 (0.029)
8 Respiratory	-0.037*** (0.006)	-0.040*** (0.008)	-0.035*** (0.008)	-0.014* (0.008)	-0.017** (0.008)	-0.038*** (0.008)	-0.035*** (0.008)	-0.020*** (0.008)
9 Digestive	-0.124*** (0.012)	-0.059*** (0.015)	-0.048*** (0.015)	-0.018 (0.014)	-0.019 (0.014)	-0.049*** (0.014)	-0.044*** (0.013)	-0.018 (0.013)
10 Genito-urinary	-0.095*** (0.016)	-0.070*** (0.021)	-0.063*** (0.020)	-0.045** (0.020)	-0.045** (0.020)	-0.065*** (0.020)	-0.054*** (0.019)	-0.038** (0.019)
11 Skin	-0.066*** (0.008)	-0.054*** (0.010)	-0.044*** (0.010)	-0.027*** (0.010)	-0.025** (0.010)	-0.046*** (0.010)	-0.033*** (0.010)	-0.017* (0.010)
12 Musculoskeletal	-0.140*** (0.005)	-0.069*** (0.007)	-0.054*** (0.006)	-0.033*** (0.006)	-0.029*** (0.006)	-0.056*** (0.006)	-0.048*** (0.006)	-0.022*** (0.006)
13 Congenital anomalies	-0.090*** (0.012)	-0.082*** (0.016)	-0.060*** (0.016)	-0.043** (0.016)	-0.035** (0.016)	-0.073*** (0.016)	-0.069*** (0.015)	-0.039*** (0.015)
14 Injuries and poisonings	-0.127*** (0.008)	-0.087*** (0.010)	-0.071*** (0.010)	-0.054*** (0.010)	-0.048*** (0.010)	-0.074*** (0.010)	-0.064*** (0.009)	-0.040*** (0.009)
Cognitive skill:	-	-	0.134*** (0.003)	-	0.110*** (0.003)	-	-	0.041*** (0.003)
Noncognitive skill:	-	-	-	0.105***	0.079***	-	-	0.054***

				(0.002)	(0.003)			(0.002)
Schooling	-	-	-	-	-	0.149*** (0.003)	-	0.052*** (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: The parameter estimates for the X variables refer to the model when diagnoses are entered by type. In Model B-H, sibling fixed effects are included. 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 *Health diagnoses*. Men born 1950-1970. Logarithm of annual earnings in 2003. Full sample.

Variable	A	B	C	D	E	F	G	H
Diabetes	-0.178*** (0.027)	-0.182*** (0.027)	-0.107*** (0.028)	-0.010 (0.028)	-0.020 (0.028)	-0.189*** (0.027)	-0.167*** (0.025)	-0.061** (0.026)
Neurosis	-0.288*** (0.008)	-0.279*** (0.008)	-0.223*** (0.008)	-0.089*** (0.008)	-0.095*** (0.008)	-0.239*** (0.008)	-0.186*** (0.008)	-0.084*** (0.007)
Personality disorder	-0.331*** (0.015)	-0.302*** (0.015)	-0.233*** (0.015)	-0.117*** (0.015)	-0.111*** (0.015)	-0.255*** (0.015)	-0.207*** (0.014)	-0.102*** (0.014)
Alcoholism and drugs	-0.409*** (0.020)	-0.381*** (0.020)	-0.317*** (0.020)	-0.180*** (0.020)	-0.183*** (0.020)	-0.310*** (0.020)	-0.234*** (0.018)	-0.120*** (0.018)
Epilepsy	-0.152*** (0.019)	-0.150*** (0.019)	-0.095*** (0.019)	-0.087*** (0.019)	-0.063*** (0.019)	-0.132*** (0.019)	-0.099*** (0.018)	-0.063*** (0.018)
Migraine	-0.034*** (0.011)	-0.033*** (0.011)	-0.040*** (0.010)	-0.010 (0.011)	-0.021** (0.010)	-0.032*** (0.011)	-0.031*** (0.010)	-0.021*** (0.010)
Asthma	-0.009 (0.006)	-0.007 (0.006)	-0.010* (0.006)	0.012* (0.006)	0.003 (0.006)	-0.009 (0.006)	-0.012** (0.006)	-0.002 (0.006)
Hay fever	0.091*** (0.004)	0.065*** (0.004)	0.036*** (0.004)	0.049*** (0.004)	0.031*** (0.004)	0.039*** (0.004)	0.032*** (0.004)	0.013*** (0.003)
Vertebrogenic pain	-0.082*** (0.005)	-0.063*** (0.005)	-0.050*** (0.005)	-0.040*** (0.005)	-0.034*** (0.005)	-0.040*** (0.005)	-0.040*** (0.004)	-0.021*** (0.004)
<i>Parental characteristics:</i>								
Father's (log) earnings	-	0.231*** (0.003)	0.178*** (0.003)	0.195*** (0.003)	0.162*** (0.003)	0.147*** (0.003)	-	0.076*** (0.002)
Father's years of schooling	-	0.011*** (0.000)	0.003*** (0.000)	0.007*** (0.000)	0.002*** (0.000)	0.000 (0.001)	-	-0.004*** (0.001)
Mother's years of schooling	-	0.012*** (0.000)	0.004*** (0.000)	0.008*** (0.000)	0.002*** (0.000)	0.000 (0.001)	-	-0.003*** (0.001)
Cognitive skill:	-	-	0.141*** (0.001)	-	0.113*** (0.001)	-	-	0.027*** (0.001)
Noncognitive skill:	-	-	-	0.132*** (0.001)	0.098*** (0.001)	-	-	0.060*** (0.001)
Schooling	-	-	-	-	-	0.173*** (0.001)	-	0.053*** (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: 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 9. The (bad) health penalty for *Health diagnoses*. Men born 1950-1970. Logarithm of annual earnings in 2003. Siblings.

Variable	A	B	C	D	E	F	G	H
Diabetes	-0.165*** (0.046)	-0.161*** (0.046)	-0.099** (0.046)	-0.043 (0.047)	-0.038** (0.047)	-0.158*** (0.045)	-0.153*** (0.042)	-0.056 (0.043)
Neurosis	-0.298*** (0.013)	-0.206*** (0.013)	-0.170*** (0.013)	-0.095*** (0.013)	-0.093*** (0.013)	-0.186*** (0.013)	-0.157*** (0.012)	-0.086*** (0.012)
Personality disorder	-0.316*** (0.023)	-0.188*** (0.022)	-0.143*** (0.022)	-0.079*** (0.022)	-0.070*** (0.022)	-0.164*** (0.022)	-0.131*** (0.021)	-0.060*** (0.021)
Alcoholism and drugs	-0.416*** (0.032)	-0.265*** (0.030)	-0.227*** (0.030)	-0.143*** (0.030)	-0.143*** (0.030)	-0.240*** (0.030)	-0.185*** (0.028)	-0.109*** (0.028)
Epilepsy	-0.140*** (0.030)	-0.137*** (0.034)	-0.091*** (0.033)	-0.090*** (0.033)	-0.066*** (0.033)	-0.130*** (0.033)	-0.112*** (0.031)	-0.074*** (0.031)
Migraine	-0.033* (0.017)	-0.002 (0.022)	-0.010 (0.021)	0.014 (0.022)	0.004 (0.021)	0.000 (0.021)	-0.008 (0.020)	-0.003 (0.020)
Asthma	0.002 (0.011)	0.000 (0.013)	-0.002 (0.013)	0.014 (0.013)	0.008 (0.013)	-0.002 (0.013)	-0.014 (0.013)	-0.007 (0.012)
Hay fever	0.087*** (0.007)	0.017** (0.009)	0.007 (0.009)	0.010 (0.009)	0.004 (0.009)	0.010 (0.009)	0.005 (0.008)	-0.003 (0.008)
Vertebrogenic pain	-0.067*** (0.008)	-0.004 (0.009)	0.000 (0.009)	0.006 (0.009)	0.008 (0.009)	0.004 (0.009)	-0.001 (0.008)	0.008 (0.008)
Cognitive skill:	-	-	0.132*** (0.003)	-	0.110*** (0.003)	-	-	0.041*** (0.003)
Noncognitive skill:	-	-	-	0.102*** (0.002)	0.076*** (0.003)	-	-	0.051*** (0.002)
Schooling	-	-	-	-	-	0.149*** (0.003)	-	0.052*** (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: The parameter estimates for the X variables refer to the model when diagnoses are entered by type. In Model B-H, sibling fixed effects are included. All regressions were run in Stata 11 using robust standard errors. Fixed effects for birth cohort were included.