SELECTION AND CAUSATION IN THE
PARENTAL EDUCATION GRADIENT IN HEALTH:
LESSONS FROM A LARGE SAMPLE OF ADOPTEES

Evelina Björkegren
Mikael Lindahl
Mårten Palme
Emilia Simeonova

Selection and Causation in the Parental Education Gradient in Health:
Lessons from a Large Sample of Adoptees
Evelina Björkegren, Mikael Lindahl, Mårten Palme, and Emilia Simeonova
NBER Working Paper No. 28214
December 2020
JEL No. I1,I12,I14,I26

ABSTRACT

We use data from a large sample of adoptees born in Sweden to study to what extent the well-established association between parental educational attainments and adult health of the child generation can be attributed to pre- or post-birth factors, respectively. We find a significant association between the educational attainment of the adopting parents and child health outcomes as adults. These results suggest that growing up in a better-educated household has long-term effects on health outcomes. Our analysis of the mechanisms behind the results suggests that formation of human capital, and in particular cognitive and non-cognitive skills, may be important.

Evelina Björkegren
Department of Economics
Stockholm University
SE-106 91 Stockholm
Sweden
evelina.bjorkegren@ne.su.se

Mikael Lindahl
University of Gothenburg
Department of Economics
Box 640
SE 405 30 Gothenburg
Sweden
and CESifo, IFAU, IZA and UCLS
Mikael.Lindahl@economics.gu.se

Mårten Palme
Department of Economics
Stockholm University
SE-106 91 Stockholm
SWEDEN
and CESifo, IFS and IZA
Marten.Palme@ne.su.se

Emilia Simeonova
JHU Carey School of Business
100 International Drive
Baltimore, MD 21202
and NBER
emilia.simeonova@gmail.com

1. Introduction

Children from families with low socioeconomic status (SES) have on average worse health outcomes than children from families with high SES. This relation has been documented in a large number of studies from different countries (see e.g. Currie, 2009, for an overview).[1] It has also been shown to hold when the health status of the children is measured late in their life and for mortality (Palme and Sandgren, 2008). It is conceivable that this health gradient in parental education to some degree is due to *causation* in the sense that it is the result of household and environmental factors associated with parental SES. However, it is not possible, based on observed associations, to rule out that the gradient is due to *selection* in the sense that genetic differences between SES groups may be transmitted between generations. To be able to say something about how much of the overall association is due to the causation, one needs to have access to data that shut down the genetic link between the parents and their children.

In this paper, we first establish a positive relationship between adult health and parental education in the population. Households with parents who are university educated have, on average, children with almost 6 percentile ranks better health as adults, compared to households with parents who only have primary education. These results are based more than 3 million individuals in the child generation. We then investigate to what extent the parental education gradient in health status is associated with pre-birth factors, primarily genetic differences, or post-birth ones, primarily environmental differences. We use a large sample of about 11,000 Swedish-born adoptees for whom we, among other things, observe educational attainments of both biological and adopting parents. If we restrict ourselves to only require identification of the biological mother, the sample size doubles. We use these samples to decompose the relationship between long-term health status and mortality and parent's educational attainment into one part due to pre-birth factors and one part due to post-birth factors.

Our results suggest that both pre- and post-birth factors matter for the relation between parental educational attainment and health outcomes for the child generation. The results from two health indices based on hospitalization data show strongly significant associations and a fairly equal split between these two factors. The results for mortality only partially confirm these results, as we find a significant association with the adoptive parents, but not with the biological parents. Comparing these estimates to those of the relationship in the population

---

[1] See Mörk, Sjögren and Svaleryd (2014) for descriptive evidence of a social gradient between child health and SES for Sweden.

suggests that our estimates obtained on adoptees are externally valid. We perform a number of sensitivity analyses, which confirm that these results are robust to issues of non-random assignment of children to adopted families and to possible post-adoption contacts between the adopted child and the biological parents.

Having shown that environmental factors associated with parents' educational attainments have lasting impact on health in the child generation, conditional on genetic and pre-natal factors, we continue by evaluating different hypotheses for mechanisms behind such a nurturing transmission. In the epidemiological literature, possible mechanisms behind health differences are often divided into two main groups: those referred to the *life course* hypothesis and those referred to the *pathway*, or *latency,* hypothesis, respectively (see e.g. Marmot *et al.,* 2001, or Case, Fertig and Paxson, 2005). The *life course* hypothesis states that environmental factors during the individuals' entire life course, including those in the very early childhood, may have separate and independent effects on health outcomes much later in life (see e.g. Almond and Currie, 2011, on long-term effects of early environmental exposure). The *pathway* hypothesis suggests that the parental education gradient is formed through different mediating factors, such as formation of skills or health-related life habits.

Our data are well suited for testing these hypotheses, since they allow us to condition on the initial health endowment of the child (through outcomes of the biological parents) and because we are able to follow the individuals over time, which includes observing many potentially important mediating factors measured in young adulthood. We investigate three different potential pathways for the parental education gradient in health. First, we look into the possibility that the effect is mediated through the children's own educational attainments. Previous studies (see e.g. Björklund, Lindahl and Plug, 2006) have shown that the educational attainments of the adopting parents are positively associated of those of their children. To the extent that there is a causal effect of education on health, which is still debated in the literature (see e.g., Galama, Lleras-Muney and von Kippersluis, 2018), the association may be attributed to the pathway of the educational attainments of the children. Second, we investigate the pathway through formation of cognitive and non-cognitive skills. Previous studies have shown a strong association between skills and health (Conti, Heckman and Urzua, 2010).

The third mediator that we analyze is formation of health-related life habits. For instance, we might think that parents with higher education transmit better dietary habits and other health-related behavior, such as lower alcohol consumption and less smoking (see e.g. Cutler and

3

Lleras-Muney, 2006, for a discussion). These habits may have direct long-lasting effect for the child, or, perhaps more likely, may form habits of the child that may be reflected in health outcomes later in life. To test for this, we use health outcomes – BMI and physical fitness – for males obtained at the time of military enlistment.

We use an econometric method for analyzing mediating factors suggested by Heckman, Pinto and Savelyev (2013), and applied to data on adoptees by Fagereng, Mogstad and Rønning (2020), to study to what extent the association between the health outcomes of the adopted children and the educational attainments of their parents is mediated by the three factors described above, and to what extent it is attributed to direct exposure. The results from this exercise unambiguously suggest that the association between adopting parents' education for child long-term health is entirely driven by the mediators, supporting the *pathway model*, and thus primarily due to investments in children's human capital and, in particular, formation of cognitive and non-cognitive skills. Our results suggest that more than 62 percent of the association can be attributed to the formation of cognitive and non-cognitive skills, around 26 percent to the educational attainment of the child, and only 12 percent to the formation of health-related behavior during early, formative years in life.

Three previous papers using data on adoptees are closely related to our study. Sacerdote (2007) study 1,650 Korean American adoptees placed by the Holt International Children's Services during 1964-1985. Showing evidence of random assignment to being placed in a highly educated and small family, relative to a large family with low education, the beneficial effects of being placed in the former family type on health-related outcomes is mixed, with positive effects for health-related behaviors such as leading to less smoking and drinking, but with no statistically significant effect on BMI (and height). Using data on foreign-born children adopted to Swedish parents, Lundborg, Nordin and Rooth (2018) show that there is a relationship between parental (especially maternal) education and youth health outcomes for sons. Osler *et al.* (2006) study the relationship between parental social class and child mortality up until age 77 using data on Danish adoptees. They show that adoptees with biological fathers from higher social class have lower mortality, but they do not find any long-term effect of adopting fathers' social class.

This paper makes several contributions *vis-à-vis* the previous literature. First, we are able to look at comprehensive measures of lifetime health based on register data on hospitalizations during almost 30 years in addition to the data on mortality. The previous studies, with the

4

exception of Osler *et al.* (2006) which looks at mortality only, use health measures obtained quite early in life. Second, again with the exception of Osler *et al.*, (2006), we are able to observe education for the biological parents of the adopted children, which can be used in order to control for pre-birth factors affecting health heterogeneity. Finally, our data obtained at different points in the life cycle of the adoptees enables us to study the mechanisms behind the results. In particular, we are able to distinguish between the *life course* and the *pathway* hypotheses for the formation of health using mediation analysis.

Our study also relates to the previous literature on the causal effects of parental education and child health. First, there is a small group of studies which have utilized exogenous changes in parental education to estimate causal effects on infant and child health, reaching different conclusions (Currie and Moretti, 2003; McCrary and Royer, 2011; Lindeboom, Llena-Nozal, and van der Klaauw, 2009; Lundborg, Nilsson and Rooth, 2014).[2] These studies estimate causal effects using arguably exogenous increases in parental education, whereas we estimate the transmission channels due to parental education and its correlates, hence capturing broader channels. The benefit from this approach is that we are able to decompose the importance of parental education into broader factors, separated into pre- and post-birth channels. Another contribution is that we apply this approach to our constructed measures of life time health. This allows us to study the long-term effects of the family educational environment.

Finally, our findings relate to the literature on intergenerational mobility in general (see e.g. Solon, 1999 and Black and Devereux, 2011, for overviews) and, in particular, the quite small literature on intergenerational persistence in health outcomes (see e.g. Andersen, 2019, Björkegren *et al.*, 2019; Halliday *et al.*, 2018; Petersen, Kragh Andersen and Sørensen, 2005; and Sørensen *et al.*, 1988). In the literature on intergenerational mobility there is an increasing number of studies using adoptees to distinguish between pre- and post-birth factors. Black *et al.* (2019) presents results for a number of outcomes, and find pre- and post-birth factors ranging from post-birth factors being much more important (e.g., for wealth) to pre-birth and post-birth factors both being important. Interestingly, for educational transmission, both pre- and post-birth factors are important and sizable (Björklund, Lindahl and Plug, 2006; Black *et al.*, 2019)

---

[2] There is also a literature that estimates the causal effect of other parental resource variables on the health of the next generation. For instance, Cesarini *et al.* (2016) find no impact on health of the next generation from exogenous positive wealth shocks for the parents through winning large sums on lotteries in Sweden. However, Akee *et al.* (2013) and Akee *et al.* (2018) find positive effects of exogenously increasing parental income on child BMI and child behavioral and emotional health in the US.

whereas for health transmission, pre-birth factors are clearly more important (see Björkegren *et al.*, 2019).

The rest of the paper is organized as follows. Section 2 presents the conceptual framework for our econometric models. Section 3 presents the data and descriptive statistics. The main results as well as sensitivity analyses are laid out in Section 4. In Section 5 we provide evidence on the mechanisms behind the findings. Section 6 concludes the paper. Finally, the paper contains an Appendix A showing results from various sensitivity analyses.

2. Empirical Specifications

2.1 Main analysis

We start by estimating the following intergenerational model on the population of non-adopted children:

$$H_i^{bc} = \beta_0 + \beta_1 Y_i^{bp} + v_i^{bc}, \tag{1}$$

where H_i^{bc} represents adult health status for the biological child and Y_i^{bp} the biological parents' educational attainment. Subscript *i* indexes the family in which the child is born and raised, and superscripts *bc* and *bp* denote the biological child and parent, respectively; v_i^{bc} is the child-specific error term assumed to be uncorrelated with Y_i^{bp}. The coefficient β_1 measures the strength of the association between adult health of the child and the educational attainment of the parents and is a combined effect of many different factors such as genetics, prenatal environment and environment during childhood and adolescence, including the causal effect of parent's education.

Using data on the characteristics of adoptees and their biological and adoptive parents, we are able to estimate the following model on the population of adoptees:[3]

$$H_i^{ac} = \alpha_0 + \alpha_1 Y_i^{bp} + \alpha_2 Y_i^{ap} + v_i^{ac}, \tag{2}$$

---

[3] We follow the strategy to separate pre- and post-birth effects from Björklund, Lindahl and Plug (2006). See also Björkegren *et al.* (2019) for a similar description as we provide in this paper.

where *Y* once again measures human capital inputs that are transmitted from the biological parent *bp*, or the adoptive parent *ap,* respectively, to the adoptive child *i*; v_i^{ac} is a child-specific error term assumed to be uncorrelated with Y_i^{bp} and Y_i^{ap}.

Using data on adopted children, and conditional on the following key assumptions of the adoption design, we are able to estimate the association between adult health status and the observable pre- and post-birth characteristics separately from equation (2): 1) Adoptees are conditionally randomly assigned to adoptive families; 2) The adoption should have taken place close to birth so that it is possible to accurately separate pre- and post-birth effects;[4] 3) The biological parents have no contact with the adopted child post adoption.[5]

Note that in general, α_2 does not only capture the importance of the adoptive parental educational attainment, Y_i^{ap}, but in addition everything else in the adoption family that is correlated with Y_i^{ap}.[6] We therefore interpret the estimates as a measure of the importance of transmission channels stemming from the pre- or post-birth influences, respectively.

Assuming that adoptees and non-adoptees are drawn from the same distribution, we are also able to decompose an estimate of β_1 into separate entities of pre- and post-birth factors, captured by estimates of α_1 and α_2, which are then interpretable for the population of children. The likelihood of generalizability of the adoption estimates increases if the intergenerational parameter is linear and if the sum of the estimates of α_1 and α_2, using the sample of adoptees, equals an estimate of β_1, obtained in the population of children. We also perform a test of the external validity of the adoption coefficients by estimating these parameters on the sample of families where at least one child has been adopted out from the family and at least one child was not adopted but is instead reared by the biological mother. We discuss these results in Section 4.2.4.

---

[4] If adoption is not taking place very shortly after birth, the postnatal pre-adoption environment (e.g., the quality of the nursery homes) needs to be uncorrelated with the genetic background and the post adoption environment (or has no influence on the health of the adopted child). Alternatively, the estimates are still unbiased estimates of the pre- and post-adoption *characteristics.*

[5] We provide a number of tests of these assumptions in Section 4.2.

[6] For a discussion about the necessary conditions under which α_2 can be interpreted as the causal effect of parents' characteristic on child characteristic, see Holmlund, Lindahl and Plug (2011).

2.2 Mediation analysis

We look at three different mediating factors as a way of studying possible mechanisms behind the results: i) education; ii) cognitive and non-cognitive skills and; iii) health behaviors, all representing traits of the child that were formed before the adult health of the child is measured. We use mediation analysis techniques for analyzing data on adoptees developed by Fagereng, Mogstad and Rønning (2020) building on work by Heckman, Pinto and Savelyev (2013) and Heckman and Pinto (2015), that provide an extension of mediation analysis to the potential outcomes framework for treatment effect estimation.[7]

The mediation analysis is separated into two parts. The first one estimates a version of equation (2) where we model the parameter α_2 to be a linear function of variables possibly mediating the relationship between Y_i^{ap} and H_i^{ac}. This leads to the following model:

$$H_i^{ac} = \delta_0 + \delta_1 Y_i^{ap} + \delta_2 M_i^{ac} + \delta_3 M_i^{ac} \cdot Y_i^{ap} + \delta_4 X_i + \delta_5 X_i \cdot Y_i^{ap} + \varepsilon_i^{ac}, \quad (3)$$

where and M_i^{ac} represent the vector of observable mediating factors, where the true mediating factors (M_i^*) can be decomposed into measured (M_i^{ac}) and unmeasured (u_i) mediating factors such as $M_i^* = M_i^{ac} + u_i$;[8] X_i represent observable individual characteristics other than Y_i^{ap} and M_i^{ac} affecting adult health outcomes, such as gender and birth cohort dummies, as well as, importantly, the educational attainment of the biological parents Y_i^{bp}.

In order to identify the importance of the mediating factors M_i^{ac}, we need two additional assumptions (compared to the analysis building on equation (2)).[9] First, we need ε_i^{ac} to be uncorrelated with M_i^{ac}. Second, we need u_i to be uncorrelated with M_i^{ac}. These are strong assumptions, since we do not have access to any quasi-experimental variation in M_i^{ac}. However, the assumptions only need to hold conditionally on observable characteristics, including the educational attainment of the biological parents of the adopted children (Y_i^{bp}). This compares

---

[7] For earlier important contributions in the econometrics of mediation analysis see Baron and Kenny (1986) and MacKinnon, Fairchild and Fritz (2007), which review key works and approaches.

[8] Equation (2), with X_i represent gender and birth cohort dummies as well as the educational attainment of the biological parents Y_i^{bp}, becomes: $H_i^{ac} = \alpha_0 + \alpha_1' X_i + \alpha_2 Y_i^{ap} + v_i^{ac}$. Let $\alpha_0 + \alpha_1' X_i = \delta_0 + \delta_2 (M_i^{ac} + u_i) + \delta_4 X_i + \mu_i$ and $\alpha_2 = \delta_1 + \delta_3 (M_i^{ac} + u_i) + \delta_5 X_i + \epsilon_i$ we get Equation (4), where $\varepsilon_i^{ac} = v_i^{ac} + \delta_2 u_i + \mu_i + \delta_3 u_i Y_i^{ap} + \epsilon_i Y_i^{ap}$.

[9] We also need ε_i^{ac} and u_i to be uncorrelated with X_i, something which we assume is true by construction.

favorably to many mediation analysis settings, which are unable to condition on any genetic background of the individuals.

If the unobserved mediating factors u_i are positively correlated with M_i^{ac}, conditional on X_i, we will understate the importance of the mediating factors. Hence, if we find these mediating factors to be important, supporting some version of the pathway model, this conclusion would likely be even stronger if these mediating factors are imperfect proxies of the true mediating factors.[10]

Finally, we estimate the relationship between each of the mediating factors M_i^{ac} and the educational attainment of the adoptive parents, conditioning on X_i:

$$M_i^{ac} = \mu_0 + \mu_1 Y_i^{ap} + \mu_2 X_i + \omega_i^{ac}. \tag{4}$$

Combining equations (3) and (4) enables us to decompose the relationship between H_i^{ac} and Y_i^{ap} into a direct and indirect effect, respectively. The latter is the part that goes through the vector of mediating factors (M_i^{ac}). For the case with only one mediator, and assuming $\delta_3 = \delta_5 = 0$, we can plug (4) into (3) to get the effect of Y_i^{ap} on H_i^{ac}, conditional on X_i, which is exactly α_2 in equation (2). This parameter α_2 is then equal to $\delta_1 + \delta_2 \mu_1$, where the former is the direct effect and the latter is the indirect effect. These components can be consistently estimated from running OLS on equations (3) and (4) conditional on the assumptions stated above.

Below, we separately investigate three possible mediators: the child's educational attainment, cognitive or non-cognitive skill and health-related behaviors. We also show results from a model where we let M_i^{ac} represent the full set of mediating variables.

3. Adoptions in Sweden, Data and Descriptive Statistics

3.1 Adoption in Sweden 1932-1967[11]

---

[10] However, one could of course also argue that the mediating factors are multidimensional in a way that makes our observed mediators negatively correlated with the unobserved mediators. This will overstate the importance of the mediating factors.

[11] A more exhaustive description of adoptions in Sweden can be found in the Online Appendix of Björkegren *et al.* (2019) posted on the home page of the Journal of Human Resources.

Several previous works describe the era of adoptions in Sweden under study in this paper. Bohman (1970) and Nordlöf (2001), focusing on the development in the Stockholm area, use primary sources. In addition, several empirical studies using data on adoptees, such as Björklund *et al.* (2004), Oskarsson *et al.* (2015) and Björkegren *et al.* (2019) give comprehensive overviews of adoptions in Sweden.

The first law regulating adoptions in Sweden was enacted in 1917. Although the law has been changed on several occasions since then, some main features of the regulation remain in place. The adoption should be finalized in a court decision and all administrative work, including the contact with the biological parents, should be done by the child welfare offices (*Barnavårdsnämnderna*). The three principles – that the adoption should be "in the best interest of the child", that no payments are allowed, and that the adopted child should have the same rights regarding inheritance as a biological child – are still applied.

Contrary to the situation today, domestic rather than international adoptions dominated. Bohman (1970) and Nordlöf (2001) give a consistent description of the mothers who gave up their children for adoption. They were on average substantially younger than the mothers who kept their children; they were, except for some exceptions, unmarried or divorced; and they had on average lower SES, although the difference to the mothers who kept their children was quite small. According to Nordlöf (2001) the largest occupational category of these mothers was maids (26 %), followed by office workers (18 %), and women employed at restaurants (15 %).

Bohman (1970) describes the biological fathers to the adopted children as on average slightly less educated, more likely to have been registered for alcohol abuse and to have been convicted compare to the fathers who did not leave their children for adoption. The predominant reason for adoption was lack of economic resources and Nordlöf (2001) attributes the sharp decline in domestic adoptions in the end of 1960s to changes in social policy, the reduced social stigma of having children out of wedlock, increased availability of contraceptives as well as legalization of abortion.

The final decision of adoption was taken by the mother after she had recovered from delivery. The child was first placed in a nursery home and thereafter placed in a prospective adoptive family. The general recommendation was that the child should be placed in the final adoptive family before six months after birth. The children underwent medical examination before they were adopted. Bohman (1970) finds no significant average differences between adopted children and children raised by their biological parents at age 10-11.

There were a few legal requirements on the adopting parents. The most important ones were that the parents should be aged as they could have been the biological parents of the child and that the adopting father should have steady employment enabling him to support the family. Until 1944 families with own biological children were not allowed to adopt. However, as documented by Nordlöf (2001), it was very rare that these families were admitted to adopt even after 1944, since it was always a shortage of children available for adoption. This convention effectively ruled out adoptions within the immediate family of the biological mothers, such as their mothers or siblings. Nordlöf (2001) estimated that such adoptions constituted less than 1 percent of the cases included in her study.

The social workers who administrated the adoptions were instructed to find a suitable homes for the children (Allmänna barnhuset, 1955). Characteristics such as height and eye color were mentioned in the instructions. However, as pointed out by Björklund *et al.* (2004), the information available on the biological parents were usually very scarce. This was also acknowledged in the instruction, which is reflected in the following quote: "The social worker's ambition to find an adoptive home that fits a specific child particularly well is often unrealistic. The important task is to find good adoptive parents who can be expected to give children in general good conditions."[12] From the instructions to the social workers there are no indications that the health status of the child was considered in any particular way in the placement.

3.2 Sample Definition

We use data from different national registers in Sweden and include all males and females born in Sweden between 1932 and 1967.[13] We use the Multigenerational Register (see Statistics Sweden, 2012) to identify whether a person was adopted as a child. This register contains personal identifier of the biological mother and father (if known to the authorities) as well as the adopting mother and father.

Table 1 shows the number of observations for the two populations used in this study – adoptees and, as a comparison, non-adoptees – at different stages of the sample selection process. In total, there are 73,384 adoptees who we can identify in our data. About 36,000 of them were adopted by only one parent, in most cases the husband of the child's biological

---

[12] Originally from Allmänna barnhuset (1969) and obtained by us from Björklund *et al.* (2004).
[13] The lower cohort restriction is motivated by data availability and the upper one by the fact that domestic adoptions in Sweden decreased rapidly in the late 1960s.

mother. We excluded these individuals from the analysis. We construct two samples from the remaining population. First, a larger sample, including 21,386 individuals for whom we have information on the biological mother as well as the adopting mother and father. Second, a smaller sample consisting of 10,880 individuals, for whom we also have information on the biological father.

Table 1. Sample sizes at different stages of the sample selection process.

Born in Sweden 1932-67	Non-adoptees	Adoptees
Non-adopted	3,711,734	
Adopted by at least one parent		73,384
Adopted by both parents		37,502
Biological mother is identified	3,541,345	24,542
Not adopted by own parents		24,498
Adopting parents' age is correct *		22,934
Not died or emigrated first year	3,529,518	22,920
Parents' education is known	3,227,149	21,386
Biological father is identified	3,227,149	10,880

*Adopting mother age 25-47 and adopting father age 25-66 at birth of adopted child.

3.3 Variable construction

3.3.1 Hospitalization in the Child generation

Data for our measures of hospitalization are obtained from the national In-patient Register (see Socialstyrelsen, 2009b). The national In-patient Register includes dates for all hospital stays at Swedish hospitals. This register has a national coverage starting in 1987, and we have access to data for the entire period until 2014. Since the first birth cohort included in our data was born in 1932, we observe all its hospital stays from age 55 and until age 82. The In-patient Register includes ICD codes for the maximum of eight different medical causes of each hospital stay.

We use two measures of health from the hospitalization data.[14] The first, labeled "Hospitalization-based health", is simply the residuals from a linear probability model regression of an indicator variable for whether or not the individual has been in hospital care for each year separately during the observation window on year and year of birth indicators. If the person is dead, we treat him or her as missing. In a second step, we average the residuals for each individual to obtain the measure. This procedure accounts for differences in the

---

[14] These are the same measures as we used in Björkegren *et al*. (2019) where we analyzed the intergenerational transmission of health and mortality.

probability of hospitalization over the life cycle and we may therefore interpret the resulting variable as a measure of lifetime hospitalization.

The second measure, labeled "Health index", is constructed in three steps.[15] First, for every year, we use a Probit model to regress an indicator variable, equal to one if the individual has died within five years and zero otherwise, on the information from the in-patient register for that year (days, visits, and diagnoses) and indicators of year of birth and gender.[16] In a second step, we create a health index ranging between 0 and 1 by predicting the risk of dying within five years. An individual is assigned the value of 1 in all years after death occurred, individuals not making any hospital visits are assigned the value of 0. Then, in a third step, we average over all years. Based on this index, we obtain a percentile rank for each birth cohort and gender separately. The difference of this measure compared to "Hospitalization-based health" is that it weights the different diagnoses by "severity" based on how likely the person is to die within five years.

3.3.2 Mortality in the Child Generation

Information on date of death is obtained from the national Cause of Death Register (see Socialstyrelsen, 2009a). The Cause of Death Register records dates and International Classification of Diseases, revision 10, (ICD 10) codes for the underlying cause of death for all deaths in Sweden from 1952 and onwards. Our observation period stops in 2016, this implies that for the child generation that we can observe the oldest person in our sample until age 84 and the youngest until age 49.

3.3.3 Educational Attainments in the Parental Generation

The number of years of schooling in the parental generation is the key independent variable in our empirical analysis. Our main data source for this variable is the 1970 Census. If the information is missing in that Census year, we use data from the 1990 and 2004 waves of the Swedish Education register. As a third option for observations that are still missing, we use the 1960 Census.[17] Overall, we are able to identify educational attainment for 97 percent of the

---

[15] The first two follow Cesarini *et al.* (2016).

[16] We use the first two digits in the ICD10 diagnosis codes (one letter and one number), which constitute about 200 different categories. We do this for the first two diagnoses for each hospital stay. In addition, we include linear variables for the number of hospital stays and the total number of days in hospital care. We control for gender and stratify on birth cohort.

[17] The education measure from 1970 is available for the population of individuals given that the individual was born in 1911 or later and was alive and lived in Sweden in 1970. It is used as the main choice because it

sample. Education in Swedish registries is recorded at seven different levels, which we translate into years of schooling.[18]

3.3.4 Mediation variables: Measures of BMI, physical fitness, cognitive, non-cognitive skills and educational attainment for the child generation

We use data on educational attainment, health and cognitive and non-cognitive skills to test the life-course and pathway hypotheses. Educational attainment is obtained from the Swedish educational registers and we use the highest level of education observed for each individual.[19] Measures of health, cognitive and non-cognitive skills are obtained for young men at age 18 at the military enlistment (army draft). The draft was compulsory in Sweden for the child cohorts included in our sample. We observe data from the army draft registry for men in the birth cohorts 1950-1967, and for these cohorts we observe information on 88% of all males in our dataset. We use two measures of health, the first is BMI and the second is a military measure of physical fitness used for placement in the military service.

The enlistment also included a cognitive skills test, which consisted of four parts testing for logical, spatial, verbal and technical abilities. We use the comprehensive measure of these different parts. An assessment of non-cognitive skills is given after a 20-30 minutes interview with a trained psychologist, with the aim of learning about the conscript's ability to cope with military service and armed combat. The conscripts are rated from 1-5 on their willingness to assume responsibility, independence, outgoing character, persistence, emotional stability, and power of initiative.[20] These sub scores are then transformed into a general measure of non-

---

measures educational attainment for individuals when they are supposed to have finished their education. Education from 1990-2004 will only be utilized for those parents that were not living in Sweden in 1970, so as not to capture educational investments later in life (which was fairly common in the 1980s and 1990s). Education in 1960, which is less detailed compared to the 1970 information, will only be used for those parents that have died before 1970 and/or that were born before 1911. A problem with the 1960 Census is that the coding of educational attainment is different from our other data sources. Therefore, we use data from individuals that are present in both the 1960 and 1970 census, and are 35-45 years old in 1960, to predict years of schooling from the 1960 census for those missing observations.

[18] Primary school (old system) or pre-comprehensive school compulsory level = 7 years; junior secondary school (old system) or comprehensive school (new system) = 9 years; vocational school =11 years, secondary school = 12 years; secondary school + 1 or 2 years = 14; college or university = 16 years; and PhD = 20 years.

[19] Educational levels are coding into years for children as we did for parents, see footnote 18. The child generation went through a reformed educational system, compared to the parental generation, leading to longer compulsory schooling. This is reflected in the years of schooling measure for the children.

[20] Other important traits in the army is the ability to cope with loss of personal freedom and socials skills, and in particular the ability to function in a group and help create group cohesion. Difficulty accepting authority and violent aggressive behavior are among the negative characteristics that psychologists look to identify. Obsession with the army and military topics is considered a negative trait. See Lindqvist and Vestman (2011), for details regarding data from the Swedish military enlistment, as well as for evidence that these measures are predictive of adult wages for the individuals.

cognitive skills. Cognitive, non-cognitive skills and physical fitness are all rated on a stanine scale 1-9, approximately following a normal distribution. We standardize these measures by year of birth, to have zero mean and standard deviation one.

3.4 Descriptive Statistics

Table 2 contains sample means and standard deviations (within parentheses) for the main outcome and control variables in the sample of adoptees and non-adoptees. Adopting fathers have almost one additional year of education as compared to the biological fathers of the adopted children. The adopting parents are on average ten years older than the biological parents of adoptees.

Although the quality of our education measure is very good, we also acknowledge that for the parental generations, there is limited variation in the lower half of the distribution. In addition to the limited variation in the actual educational outcomes in these birth cohorts, for those born before 1911 we are restricted to use data from the 1960's census, with less detailed information on educational outcomes compared to later censuses. This is especially problematic when we analyze mortality of the child, since the variation of this outcome primarily comes from those children born in the first cohorts included in the sample. In our main analysis we therefore use a measure of parental education calculated as the average of years of schooling for mothers and fathers, hence maximizing the variation in the lower end of the distribution of the parental education variable.

Table 2. Summary statistics of main outcome and control variables

	Non-adoptees			Adoptees		
	Mean	Std. Dev	N	Mean	Std. Dev	N
Children, all						
Female	0.49	(0.50)	3,227,149	0.48	(0.50)	10,880
Year of birth	1951.78	(9.58)	3,227,149	1956.80	(6.70)	10,880
Share of dead	0.12	(0.32)	3,227,149	0.10	(0.30)	10,880
Hospitalization	50.03	(28.20)	3,199,618	44.64	(29.02)	10,845
Health index	50.03	(28.15)	3,227,149	45.08	(28.58)	10,880
Education	11.47	(2.69)	3,204,124	11.52	(2.18)	10,812
Children, men						
Non-cognitive	0.00	(1.00)	794,795	-0.09	(1.03)	4,013
Cognitive	0.00	(1.00)	803,957	-0.18	(0.98)	4,061
Physical fitness	0.00	(1.00)	802,792	-0.09	(0.98)	4,050
BMI	21.47	(2.82)	794,722	21.60	(3.02)	4,018
Biological mother						
Year of birth	1923.58	(11.92)	3,227,149	1932.74	(8.89)	10,880
Education	8.18	(2.16)	3,227,149	8.18	(1.92)	10,880
Biological father						
Year of birth	1920.06	(12.49)	3,227,149	1929.14	(9.87)	10,880
Education	8.69	(2.77)	3,227,149	8.43	(2.27)	10,880
Biological parents						
Education	8.44	(2.17)	3,227,149	8.31	(1.64)	10,880
Adopting mother						
Year of birth				1922.61	(8.85)	10,880
Education				8.62	(2.51)	10,880
Adopting father						
Year of birth				1919.87	(9.23)	10,880
Education				9.38	(3.12)	10,880
Adopting parents						
Education				9.00	(2.49)	10,880

Note: Means and standard deviations of main control and outcome variables.

3.5 The Association between Biological and Adopting Parent Characteristics

A possible concern with the interpretation of the coefficient estimates is that of selective placement of adoptees. Table 3 illustrates the correlation in years of education between adopting and biological parents of adoptees.

Table 3. Correlation between biological and adoptive parents' education

	Mothers	Fathers
Educational attainment	0.1572	0.1708

Note: Education is standardized by birth cohort. *p*-values are below 0.1 percent.

The correlation for years of schooling is quite similar to that reported by Björklund, Lindahl and Plug (2006) for children in cohorts born 1962-1966. There are at least two reasons why we would observe a positive correlation for characteristics of biological and adoptive parents. First,

this could happen if some adoptions are made by relatives of one of the biological parents. Second, there could be matching on characteristics known to the adoption agency, either because of the demand of parents, or because of a view that an adopted child would be better off in an adoptive family with similar characteristics as the biological parents.

As discussed in Section 4.2, the empirical importance of the first reason – adoptions by relatives – is likely to be very limited since the rule of not allowing people with own biological children to adopt to a large extent precluded parents and siblings of the biological parents from doing that. Nordlöf (2001) estimated these adoptions to be around 1 percent of the total number of adoptions in the Stockholm area. Brandén, Lindahl and Öckert (2018) confirm this conclusion, although their estimate of the share of adoptions by close relatives is slightly higher at 5.4 percent, applying to the whole country. They are also able to eliminate those adopted by close relatives from their sample and find that the correlation in years of schooling between (unrelated) adoptive and biological parents of adoptees remains virtually unchanged.

The second reason, matching, is likely to be a more important mechanism. If this matching is made on characteristics observable in the data (such as educational attainment or health characteristics), we are able to control for this in the estimations. In a sensitivity analysis, we will investigate this further by including more detailed health and education data of the biological (adoptive) parents and see what happens to the estimate for the characteristics of adoptive (biological) parents (a similar test was made in Björklund *et al.* 2006). If we do not see any change, which is what happens (see Section 4.2), we can rule out matching on observable characteristics as affecting our conclusions.

The remaining reason would then be matching on characteristics that are unobservable in the data. Björklund *el al.* (2006) investigate this issue by deriving the magnitude of the bias (modeled as a combination of selective placement and measurement error), finding evidence that the bias accounts for at most 13 percent of the estimated impact of the adoptive and biological parents' characteristics on adoptees' educational attainment. We return to and further examine this issue in Section 4.2.

A second potential threat to the random assignment assumption is that adoptees may be non-randomly assigned to adoptive families based on health endowments at birth. This would be particularly troubling if more educated adoptive parents are able to select healthier children. While we cannot directly test for this because we lack data on health at birth, based on the

previous literature, we believe this is unlikely to happen.[21] Nevertheless, in Section 4.2, we perform sensitivity analysis as to whether selective placement is likely to affect our results, concluding that this is unlikely to be the case.

4. Results

4.1 Health Measures Based on Hospitalization Data

Figure 1 shows the relation between parental education, measured as the average years of schooling of the parents, and adult health status, measured by our two hospitalization indices, both in percentile ranks, where higher ranks indicate better health. We plot the average rank against the parental education categories, which are based on the average years of schooling for the parents.

As expected, there is a visible pattern of increasing health by parental education in the samples of non-adoptees (the two top panels a and b). We see that households with parents who are at the top of the education distribution have, on average, children with 10-15 percentile ranks better health, compared to households with parents who are at the bottom of the education distribution. The associations are well approximated by a linear relationship. [22]

The lower four panels, c through f, plot the relation between parental education and adult health status for adoptees, where the latter is either related to biological parents (panels c and d) or adoptive parents (panels e and f). The positive relationship applies to both biological and adopting parents in the adoptee sample. Again, the associations are fairly well approximated by a linear relationship, except at the top of the educational distribution.[23]

---

[21] As argued in Björkegren *et al.* (2019): First, the institutional set up at the time was such that adoptive families were approached as soon as a candidate for adoption became available and there was an excess of candidate adoptive parents relative to available children. Second, unhealthy infants that were given away by their biological mothers were not offered for adoption. Finally, Holmlund, Lindahl and Plug (2008) show that there is no significant correlation between adoptive parents' education, the gender of the adoptee and the biological mother's age at birth – the only two pre-existing characteristics that are available in the data and could potentially proxy for infant health at birth

[22] Mörk, Sjögren and Svaleryd (2014) provide similar evidence for child health, finding that children in families with the lowest SES have 10-15 percentile higher risk of being hospitalized compared to children in families with the highest SES in Sweden.

[23] There is a difference in the share with high education among biological and adopting parents. Among the adopted children's biological mothers, 4 % have more than 11 years of education and 2.4 % have 14 years or more. Among adopting mothers 10 % have more than 11 years of education, and 7.9 % have 14 years or more. Thus, the precision of the estimates for this highest educated group of mothers across the biological and adoptive mothers' samples varies significantly.
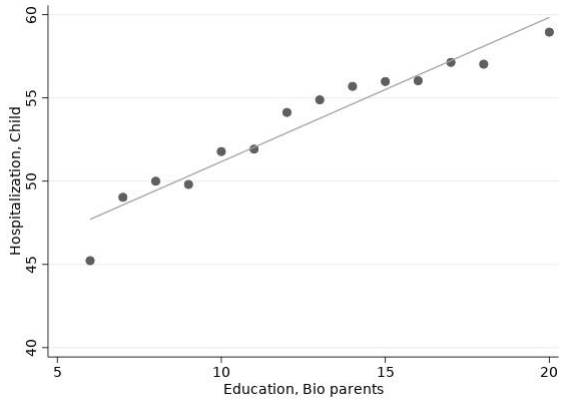
a) Hospitalization, non-adoptees
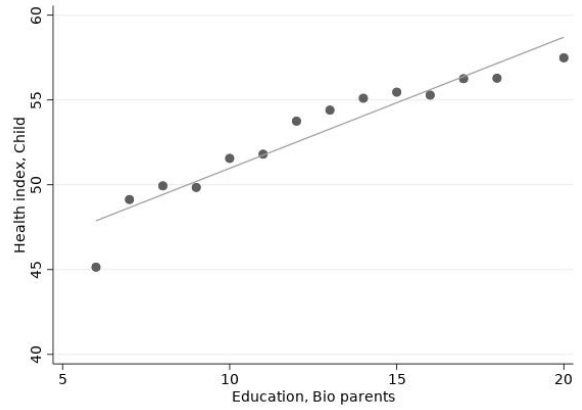
b) Health index, non-adoptees
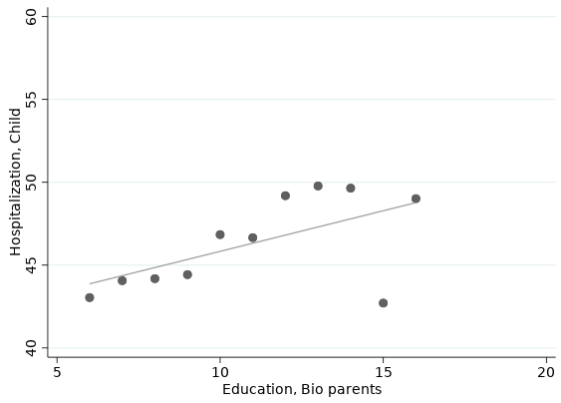
c) Hospitalization, adoptees and biological parents
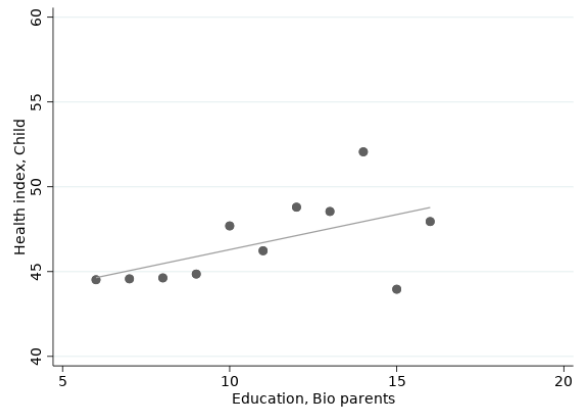
d) Health index, adoptees and biological parents
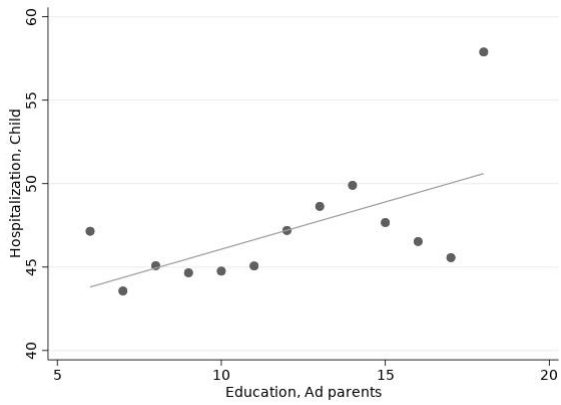
e) Hospitalization, adoptees and adopting parents
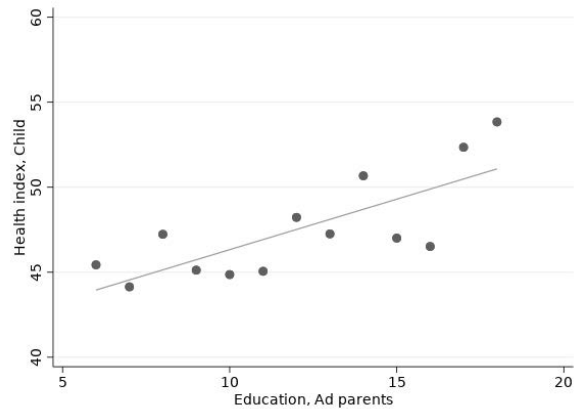
f) Health index, adoptees and adopting parents

Note: The figures plot the relationship between rank child health indices and average parental education (in years) for non-adoptees and adoptees. Because of small samples in some educational groups, we round off average educational years to the closest integer. Categories containing less than 10 observations are grouped with the closest category with at least 10 observations.

Figure 1. Mean percentile rank of adult health in the child generation and average parental education

Table 4 shows the results from our linear regression model for the association between parental educational attainments and the ranks of hospitalization (the first two columns) and health index (the last two columns) in the child generation. The results for non-adoptees, shown in columns (1) and (3), reveal a highly significant association between health and the educational attainments of the parents. One more year of schooling of both parents is associated with almost one percentile rank better health of their children when they are adults or of old age. This means that households with parents who are university educated have, on average, children with almost 6 percentile ranks better health as adults, compared to households with parents who only have primary education. In standard deviation units, this is equivalent to one standard deviation higher parental education being associated with about 0.07 standard deviation better health for the children as adults.

A decomposition of the relative influence of the biological and adopting parents, respectively, gives a close to 50/50 split between biological and adopting parents for the hospitalization measure (column 1 and 2). The estimates for the health index (column 3 and 4) are similar to the estimates for hospitalization and confirms that adopting and biological parents' education are equally important for long-term health.

Table 4. Associations between health indices (in percentile rank) and parental education (in years).

	(1)	(2)	(3)	(4)
	Hospitalization		Health index	
	Non-adoptees	Adoptees	Non-adoptees	Adoptees
Education, Bio parents	0.883***	0.507***	0.801***	0.420**
	(0.007)	(0.180)	(0.007)	(0.177)
Education, Ad parents		0.488***		0.352***
		(0.120)		(0.119)
N	3,199,618	10,845	3,227,149	10,880

Note: Results from OLS regressions. Standard errors in parentheses; *** significant at 1%, ** at 5%, * at 10%. Each column represents a separate regression and all regressions include indicators for gender and birth cohort of children, and five-year intervals for parental cohorts. Column (1) and (3) is based on a sample of non-adopted children, column (2) and (4) on adoptees for whom we have information on all parents.

Appendix Table A1 shows results for mothers and fathers, separately. It also shows results from the larger sample of adoptees, where we also include adoptees with unknown biological fathers.

4.2. Sensitivity Analyses

4.2.1 Tests for selective placement of adoptees to families

To test whether adoptees are conditionally randomly assigned to adoptive families, we do three sets of sensitivity analyses. First, we look at the robustness with respect to changes in the set of confounding parental characteristics included in the model. In Section 3.4, we showed that there is a positive correlation between biological and adopting parents' education. By always including biological parents' education in our regressions, we mitigate the main concern that the estimate for adopting parents are capturing the importance of biological parent's education. We might, however, still worry that there are unobserved factors that are correlated and that ignoring these could bias the estimates.[24]

In Appendix Table A4, we show the robustness of our main results for hospitalization when we first only include information on adoptive parents, and then add relevant possible controls for biological parents, first only education and cohort controls (which is the specification used in column 4 of our main results Table 4) and then additional controls for the health (and location) of the biological parents. The purpose is to study how sensitive the coefficient for adoptive parents is due to omitting variables of the biological family, especially their health, which we know is predictive of the adopted child's health (see Björkegren et al., 2019). Although we do see that the adoptive parent's education coefficient is affected by inclusion to the added parental health variables it remains statistically significant also with the additional controls included in the specification. We then reverse the order and sequentially add characteristics of the adoptive parents to investigate how sensitive the coefficient for biological parents is for omitting variables of the adoptive family. Again, the results in Appendix Table A4 are shown to be very robust.

Second, we cannot directly observe whether relatives or friends of the biological parents adopted some children, but in such cases, children are more likely to stay in the municipality where they were born. Moreover, adopted children who move from their municipality of birth are much less likely to interact with their biological parents post adoption. To address the potential problem resulting from the risk that biological parents might have contact with the child post-adoption or that there might be closer matching of parents locally, we restrict the

---

[24] Previous research using these data has shown that the resulting bias coming from matching of biological and adopting parents' characteristics is likely to be small. See for example Black *et al.* (2019), Björkegren *et al.* (2019) and Björklund, Lindahl and Plug (2004).

sample to children who were adopted to families living in another municipality.[25] Appendix Table A5 shows that the main results remain very similar in this sample, confirming that the potential bias likely to be small.

Third, for the subsample of biological mothers who have given birth to more than one child, and where at least two of the children has grown up in other families, we regress health on the adopting mother's education, controlling for biological siblings fixed effects.[26] This means that we control for selective placement due to *unobservable* characteristics stemming from the biological mother. Results are shown in Appendix Table A6, and, although imprecisely estimated, the estimates for the adoptive mother including fixed effects (in columns 2 and 4), are very similar to the one only controlling for observable biological characteristics (in columns 1 and 3). Overall, we conclude that selective placement is unlikely to affect our main conclusions.

4.2.2 External validity and nature-nurture interactions

Adoptees in our sample are born in families that on average have lower SES than the population in general. To investigate to what extent the results have external validity, we study a sample of children to biological mothers who have given birth to at least two children - raised at least one herself and given up at least one for adoption.

The results in Table 5 show that there is a significantly stronger correlation between health and parental educational attainment in this sample. This result suggests that there might be some non-linearities in the relationship between parental education and health of the next generation.

---

[25] We make use the Census in 1960 where we have information on the municipality where both adopting and biological mothers live.

[26] See Björklund, Lindahl and Plug (2006).

Table 5. Associations between health indices (in percentile rank) and parental education (in years). External validity. Sample of children born to mothers who gave up at least one child for adoption and kept at least one child.

	(1)	(2)	(3)	(4)
	Hospitalization		Health index	
	Non-adoptees	Adoptees	Non-adoptees	Adoptees
Education, Bio parents	1.272***	0.413*	1.271***	0.373*
	(0.103)	(0.214)	(0.102)	(0.210)
Education, Ad parents		0.475***		0.320**
		(0.143)		(0.141)
N	28,516	8,160	28,817	8,184

Note: Results from OLS regressions. Standard errors in parentheses; *** significant at 1%, ** at 5%, * at 10%. Each column represents a separate regression and all regressions include indicators for gender and birth cohort of children, and five-year intervals for parental cohorts. The sample only include children with biological mothers that gave up at least one child for adoption and kept at least one child.

So far, we have estimated separate effects of biological and adopting parents' education on health in the child generation. It is however possible that there might be important interaction effects between these two factors that we should consider to better understand how health capital is formed. However, the results from a model that includes interactions effects between the educational attainment of biological and adopting parents have produces too imprecise estimates to draw any conclusions regarding existence of such interaction (see Appendix Table A7).

4.3 Mortality

Table 6 shows hazard ratio estimates for the association between mortality and parental educational attainments. For the sake of comparison, column 1 shows the results for children raised by their biological parents. The results in this sample suggest that one year additional schooling in the sum of parental schooling is associated with a decrease in mortality in the child generation by about 5 percent. The results shown in column 2 reveal that this association can be fully attributed to post-birth influences. An extra year of education for the sum of the adopting parents' schooling is associated with a 4 percent decrease in mortality, while there is no significant association with the educational attainments of the biological parents.

Table 6. Associations between mortality and parental years of schooling (Hazard ratios).

	(1) Non-adoptees	(2) Adoptees
Education, Bio parents	0.948***	1.032
	(0.001)	(0.021)
Education, Ad parents		0.960***
		(0.015)
N	3,227,149	10,880

Note: Results from Cox proportional hazard models. Standard errors in parentheses; *** significant at 1%, ** at 5%, * at 10%. Each column represents a separate regression and all regressions include indicators for gender and birth cohort of children, and five-year intervals for parental cohorts. Column (1) is based on a sample of non-adopted children, and column (2) a sample of adoptees.

Appendix Table A8 shows results separately for mothers and fathers and also adds results from the extended sample also including adopted children with unknown biological fathers. The results on the population of children raised by their biological parents show that the protective effect of parental education are fairly equally shared between the educational attainments of mothers and fathers. The results on the sample of adoptees show a marginally significant effect of the adopting mothers' years of schooling on child mortality. Finally, Appendix Table A8 show that the results obtained on the extended sample are very similar to the ones for the sample including information on both biological parents.

Appendix Table A9 shows results for different causes of death. These results indicate that being adopted by parents with higher education reduces the risk of dying from cancer as well as preventable diseases. As defined in Appendix Table A10, the causes of death defined as *preventable diseases* are mainly conditions related to smoking and alcohol consumption.

5. Mechanisms: Results from the Mediation Analysis

Table 7 presents the results from the estimates of equations (3) and (4) in Section 2.2.[27] We use *hospitalization* as the health outcome variable for the child generation, since this variable provides more precise estimates than the ones obtained from the *health index*.[28] Column 1 shows parameter estimates from an OLS regression of equation (4) for the sample of men in the child generation for which we have access to the mediation variables *Educational attainment*, *Cognitive skills*, *Non-cognitive skills*, *Physical fitness* and *BMI*. The latter four measures are obtained from the military conscription at age 18-19 (see Section 3 for details).

---

[27] We here confine the presentation of the results to where we impose the restriction $\delta_3 = \delta_5 = 0$ in order to save space.

[28] The results for the health index are similar to those in Table 7, but less precise (see Appendix Table A11).

Columns 2-5 present results from various versions of Equation (3) where we have added these five mediation variables to the model. Finally, columns 6-10 provide results from equation (4), using the mediation variables as outcomes.[29]

As explained in Section 2.2, the overall role of the mediation variables can be calculated in two different ways. First, as the product of the vector of the coefficient estimates of the mediation variables for one of the models estimated in columns 2-5, and the corresponding coefficient estimates in columns 6-10. Second, as the difference between the coefficient estimates for adoptive parent's education in column 1 and the corresponding coefficient estimate in columns 2-5.

It is evident from the estimates for adoptive parent's educational attainment in column 5 that the mediation variables explain the full effect of the adoptive parents' education on health in the child generation. This lends strong support for the *Pathway hypothesis* for the health-parental SES association, i.e. that the primary effect of parental education is on educational attainments of the children, which, in turn, affects their health outcomes as adults, as opposed to the *Life course hypothesis*, which asserts a separate, and lasting, effect of environmental factors during childhood on health outcomes much later in life.

Next, we decompose the mediation effects into the contribution of the separate mediation variables. We use the estimates from the specification in column 5 of Table 7 including all mediating factors. The results are summarized in Figure 2. The figure illustrates four sets of results with different coefficient restrictions. The lowest bar represents the results corresponding to Table 7, under the restriction $\delta_3 = \delta_5 = 0$. We find that the human capital variables play the dominant role in formation of health: 26 percent can be attributed to educational attainments and 62 percent to the skills variables, where non-cognitive skills are the largest contributor.[30] The latter result is especially noteworthy since the skills variables are measured at age 18-19, before many young people in the cohorts under study have finished their education.

The two top bars present results when we relax the two restrictions one at a time, and the third bar presents the results for the most flexible model, without any restrictions on coefficients in equation (3). The main message is the same in all four sets of results: health mediates between

---

[29] Lundborg, Nordin and Rooth (2018) and Brandén, Lindahl and Öckert (2018) have previously analyzed the impact of parents' educational attainment on variables from the military enlistments data using foreign adoptees.
[30] The separate contributions can be calculated as the product of the estimate for the respective mediator in column 5, times the estimate on adoptive parent's education for the respective outcome in columns 6-10.

6 and 12 percent, education between 24 and 26, and skills are the most important factor in all three, mediating between 61 and 66 percent.

The strong role for cognitive and non-cognitive skills is in line with previous research on the associations between these outcomes and health outcomes (see e.g. Batty and Deary, 2004, or Osler *et al, 2003,* on cognitive skills, and Smithers *et al.*, 2018, on non-cognitive skills). Recent research that emphasizes social interactions and networks, associated with non-cognitive skills, as being key factors for remaining in good health as the individuals ages (see e.g. Smith and Christakis, 2008). We find very limited evidence of a role for the health variables: *Physical fitness* and *BMI* measured at age 18. We interpret this as suggestive evidence that the pathway hypothesis mainly works via building the stock of human capital, but not via the stock of health capital.

Since the estimates in Table 7 are based on conscription data, which is only available for men, and for a certain period, the estimates are fairly imprecise. To assess to what extent our main conclusions hold up in a larger sample, also including women and men in all birth cohorts, we perform a mediation analysis including educational attainment of the child as possible mediator, only.

Table 8 shows the results equivalent to the ones in columns 1, 5 and 8 in Table 7.[31] These estimates show that educational attainment of the child explains about 31 percent, unconditionally on other possible mediators. When we calculate this number based on the estimates in columns 1, 5 and 8, of Table 7, we find education to explain about 65%, unconditional on the other mediators. The difference appears to be mainly due to a difference between education as a mediating channel for men and for women.

The mediation analysis here has focused on analyzing the role of the child variables as mediators between adult health of the child and the *adoptive* parents' educational attainment. These child variables can, however, also act as mediators between adult health of the child and the *biological* parents' educational attainment. This would happen if, for example, inherited genetic factors relevant for human capital creation affect child adult health through children's human and health capital formation. If we look at the results in Table 7, we see that, although the mediators then play a less important role, they explain roughly 60 percent of the link between biological parents' educational attainment and child adult health (see columns 1 and 5

---

[31] The results for Health Index are shown in Table A12.

in Table 7).

**Table 7.** Mediation analysis – Hospitalization as health outcome in adulthood

	(1)	(2)	(3)	(4)	(5)	(6) BMI	(7) Physical fitness	(8) Education	(9) Non-cognitive	(10) Cognitive
			Hospitalization							
Education, Ad parents	0.340*	0.253	0.120	0.033	-0.090	-0.058**	0.020***	0.138***	0.049***	0.071***
	(0.190)	(0.190)	(0.193)	(0.192)	(0.192)	(0.024)	(0.007)	(0.014)	(0.007)	(0.006)
Education, Bio parents	0.944***	0.864***	0.674**	0.516*	0.419	-0.023	0.037***	0.169***	0.064***	0.102***
	(0.287)	(0.285)	(0.288)	(0.289)	(0.287)	(0.031)	(0.010)	(0.022)	(0.010)	(0.010)
BMI		-0.954***			-0.926***					
		(0.167)			(0.162)					
Physical fitness		1.548***			-0.078					
		(0.482)			(0.515)					
Education			1.595***		0.807***					
			(0.229)		(0.253)					
Non-cognitive				2.415***	2.462***					
				(0.489)	(0.523)					
Cognitive				2.678***	2.078***					
				(0.518)	(0.549)					
N	3,942	3,942	3,942	3,942	3,942	3,942	3,942	3,942	3,942	3,942

Note: Results from OLS regressions. Standard errors in parentheses; *** significant at 1%, ** at 5%, * at 10%. Each column represents a separate regression and all regressions include indicators for gender and birth cohort of children, and five-year intervals for parental cohorts. The sample is based on male adoptees with information on educational attainment and measures on health and skills from military conscription.

Table 8. Mediation analysis: Hospitalization as health outcome in adulthood. Both female and male adoptees, focusing at child's education as mediating variable.

	(1) Hospitalization	(2) Hospitalization	(3) Education
Education, Ad parents	0.495***	0.333***	0.152***
	(0.120)	(0.122)	(0.009)
Education, Bio parents	0.477***	0.281	0.184***
	(0.180)	(0.182)	(0.014)
Education		1.062***	
		(0.134)	
N	10,791	10,791	10,791

Note: Results from OLS regressions. Standard errors in parentheses; *** significant at 1%, ** at 5%, * at 10%. Each column represents a separate regression and all regressions include indicators for gender and birth cohort of children, and five-year intervals for parental cohorts. The sample is based on all adoptees with information on educational attainment.
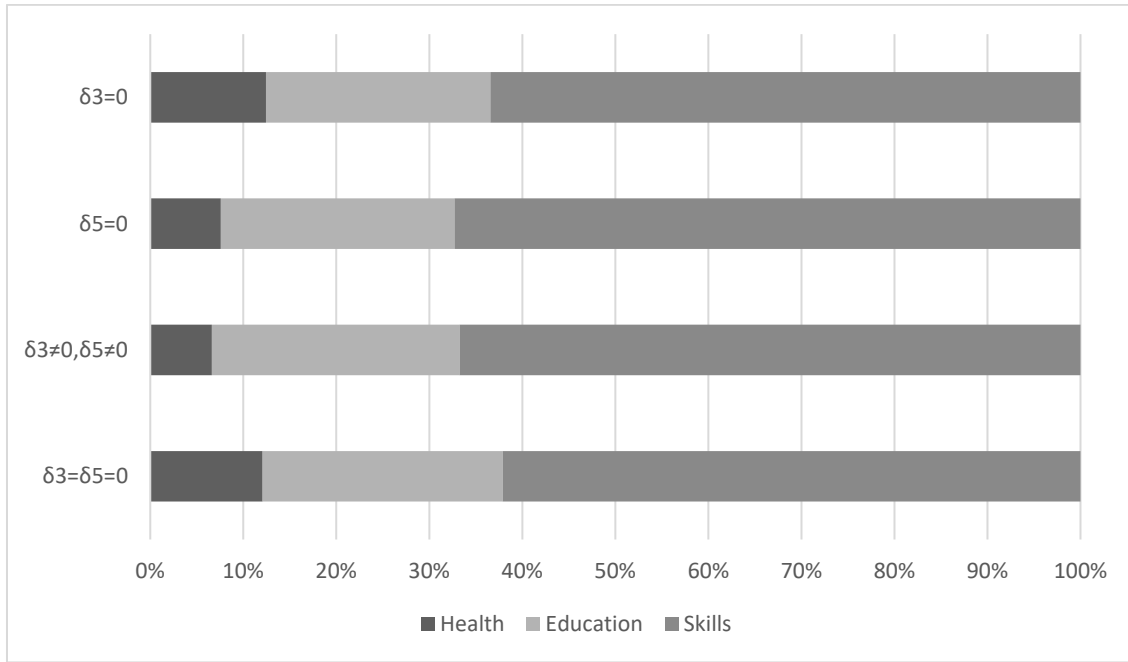
Note: This figure decomposes the association between adopting parents' education and child long-term health by attributing it to three key groups of observable mediator variables, as described in Section 2.2. We report results with four different sets of restrictions on coefficients in equation (3).

Figure 2. Decomposition of indirect link between adult health and adopting parent's education, which is attributed to key groups of mediator variables.

6. Conclusions

In this study we use a large sample of adoptees, born between 1932 and 1967, and decompose the parental education gradient in child long-term health into one component reflecting the influence of pre-birth characteristics (*selection*) and one component reflecting post-birth influences of parental educational attainments as well as unobservable characteristics associated with it (*causation*). In a second part of the study we look into the mechanism behind primarily the second component of the decomposition using a mediation analysis technique (see e.g. Heckman *el al.,* 2013, or Fagereng *et al.,* 2020).

Three main conclusions can be drawn from our results. First, the parental education gradient can be attributed to both pre- and post-birth factors. Both components are significantly different from zero when we use the *Hospitalization* as well as the *Health index* as outcome measures of long-term health in the child generation. Second, we can reject that educational attainments of the adopting parents is not associated with all three main health outcomes included in the study, i.e., this applies also for the measure of life expectance. Third, the

28

mediation analysis suggests that the association between health and parents' educational attainments can primarily be attributed to improved build-up of human capital, including cognitive and non-cognitive skills in homes with more educated parents.

In the introduction, we highlighted three previous studies of particular relevance to the present one. How do our results relate to the ones obtained in these papers? First, Osler *et al.* (2006) shows a significant association between the SES level of the biological father and mortality in the child generation, but no significant associations to the SES of the adopting father, i.e., different from what we find on mortality. However, bearing in mind the follow up period on mortality, by necessity, is limited, to age 77 in the Osler *et al.* study and 83 for the oldest cohort included in this study, it is inevitably the case that censoring of the outcome severely restricts the precision of the estimates. In the main estimates of the Osler *et al.* study, the lower limit of the 95 percent confidence interval is a 16 percent lower mortality of being placed in a home with a high/middle versus working class. This underlines the importance of using health measures based on mid-life outcomes as a complement to mortality in intergenerational studies, where mortality is most often heavily censored for at least one of the generations.

The results of Lundborg *et al.* (2018) obtained on Swedish conscription data –marginally significant associations between parental educational attainments and youth health outcomes as well as strong associations to cognitive and non-cognitive skills as well as child education – concurs, reassuringly, with the results presented in Table 7 columns 6 through 10 in this paper, although slightly different samples are used.[32] Finally, Sacerdote (2007) finds association between being placed in a small high educated family in a sample of Korean adoptees in the US on early health outcomes and health related behavior, such as BMI, obesity, drinking and smoking behavior. These results support our general findings, and in particular, the association between being placed in a small educated family and smoking obtained by Sacerdote can serve as a possible background to our result on adopting parent education and mortality in preventable diseases.

The intergenerational persistence in socio-economic and health outcomes as well as inequalities in opportunities have largely been overlooked in the previous research on the causes behind the education gradient in health (see Galama *et al*., 2018, for a recent overview). Although this paper is an attempt to fill parts of this gap there is obvious need for further

---

[32] Lundborg *et al.* (2018) use foreign born adoptees born between 1965 and 1978.

research in this area. The most immediate direction for future research concerns evaluations of public policy interventions – both in the health and education sectors – aiming at promoting equality of opportunities and counteracting the education gradient in health.

References

Akee, R., E. Simeonova, W. Copeland, A. Angold and J. Costello (2013) "Young Adult Obesity and Household Income: Effects of Unconditional Cash Transfers" *American Economic Journal: Applied Economics*, 5(2), 1-28.

Akee, R., E. Simeonova, W. Copeland and J. Costello (2018) "Can Household Income Affect Child Personality Traits and Behaviors?" *American Economics Review*, Vol. 108, No 3, March 2018.

Allmänna barnhuset (1955) *Adoption. Handbook for social workers and authorities responsible for adoption issues*. Stockholm.

Allmänna barnhuset (1969) *Adoption. Handbook for social workers and authorities responsible for adoption issues*. Stockholm.

Almond, D. and J. Currie (2011) "Human capital development before age five" in David Card and Orley Ashenfelter (eds.) *Handbook of Labor Economics*, *4*, 1315-1486.

Andersen, Carsten (2019) "Intergenerational Health Mobility: Evidence from Danish Registers" Economics Working Papers 2019-04, Department of Economics and Business Economics, Aarhus University.

Baron, R. M., D. A. Kenny (1986) "Moderator-Mediator Variables Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations." *Journal of Personality and Social Psychology* 51(6), 1173-82.

Batty, G. D. and Deary, I. J. (2004). "Early life intelligence and adult health." *British Medical Journal* 329, 585-586.

Björkegren, E., M. Lindahl, M. Palme and E. Simeonova (2019) Pre-and Post-Birth Components of Intergenerational Persistence in Health and Longevity Lessons from a Large Sample of Adoptees, *Journal of Human Resources*, forthcoming.

Björklund, A., M.Lindahl, and E. Plug (2004) "Intergenerational Effects in Sweden: What Can We Learn from Adoption Data?" IZA, Discussion Paper 1194.

Björklund, A., M. Lindahl, and E. Plug (2006) "The Origins of Intergenerational Associations: Lessons from Swedish adoption data" *The Quarterly Journal of Economics*, 999-1028.

Black, S. E. and P. J. Devereux (2011) "Recent Developments in the Study of Intergenerational Mobility" in David Card and Orley Ashenfelter (eds.) *Handbook of Labor Economics*, 4, 1487-1541.

Black, S.E., P. J. Devereux, P. Lundborg, and K. Majlesi (2019) Poor Little Rich Kids? The Role of Nature versus Nurture in Wealth and Other Economic Outcomes and Behaviors. *Review of Economic Studies*, forthcoming.

Bohman, M. (1970) *Adopted Children and their Families.* Stockholm: Proprius.

Brandén, G., M. Lindahl and B. Öckert (2018) "The importance of nature-nurture interactions for socio economic outcomes: Evidence from a large sample of adoptees", Mimeo, University of Gothenburg.

Case, A., A. Fertig and C. Paxson (2005). "The lasting impact of childhood health and circumstance" *Journal of Health Economics*, *24*, 365-389.9

Cesarini, D., Lindqvist, E., Östling, R., & Wallace, B. (2016). Wealth, health, and child development: Evidence from administrative data on Swedish lottery players. *The Quarterly Journal of Economics*, *131*(2), 687-738.

Conti, G., Heckman, J., & Urzua, S. (2010). The education-health gradient. *American Economic Review*, *100*(2), 234-38.

Currie, J (2009) "Healthy, Wealthy and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development" *Journal of Economic Literature*, Vol 47, No 1, pp 87-122

Currie, J., and E. Moretti. (2003) "Mother's education and the intergenerational transmission of human capital: evidence from college openings" *The Quarterly Journal of Economics*, 1495-1532..

Cutler, D. M. and A. Lleras-Muney (2006) "*Education and health: evaluating theories and evidence*" National Bureau of Economic Research, Working Paper w12352.

Fagereng, A. M. Mogstad and M. Rønning (2020) "Why do wealthy parents have wealthy children?", *Journal of Political Economy,* forthcoming.

Galama, T., A. Lleras-Muney and H van Kippersluis (2018) "The Effect of Education on Health and Mortality: A review of the Experimental and Quasi-Experimental Evidence," *Oxford Research Encyclopedia of Economics and Finance*, The Oxford University Press.

Halliday, T., B. Mazumder and A. Wong (2018). "Intergenerational Health Mobility in the US" FRB of Chicago, Working Paper No. 2018-2.

Heckman, J. J. and R. Pinto (2015). "Econometric mediation analyses: Identifying the sources of treatment effects from experimentally estimated production technologies with unmeasured and mismeasured inputs." *Econometric Reviews* 34 (1-2), 6–31.

Heckman, J. J., R. Pinto, and P. Savelyev (2013). "Understanding the mechanisms through which an influential early childhood program boosted adult outcomes." *American Economic Review* 103 (6), 2052–2086.

Holmlund, H., M. Lindahl, and E. Plug (2008) "The causal effect of parents' schooling on children's schooling: A comparison of estimation methods" *IZA Working Paper 3630, August* 2008

Holmlund, H., M. Lindahl, and E. Plug (2011) "The causal effect of parents' schooling on children's schooling: A comparison of estimation methods" *Journal of Economic Literature* 49(3), 615-651.

Lindeboom, M., A. Llena-Nozal, and B. van Der Klaauw. (2009) "Parental education and child health: Evidence from a schooling reform." *Journal of Health Economics* 28(1), 109-131.

Lindqvist, E., and Vestman, R. (2011). The labor market returns to cognitive and noncognitive ability: Evidence from the Swedish enlistment. *American Economic Journal: Applied Economics*, *3*(1), 101-28.

Lundborg, P., A. Nilsson and D. O. Rooth (2014) "Parental education and offspring outcomes: evidence from the Swedish compulsory School Reform" *American Economic Journal: Applied Economics 6*(1), 253-278.

Lundborg, P., M. Nordin and D. O. Rooth (2018) "The Intergenerational Transmission of Human Capital: the Role of Skills and Health" *Journal of Population Economics* 31(4), 1035-1065.

MacKinnon, D., A. Fairchild, and M. Fritz (2007) "Mediation Analysis" *Annual Review of Psychology* 58, 593-614

Marmot, M., M. Shipley, E. Brunner, and H. Hemingway (2001) "Relative contribution of early life and adult socioeconomic factors to adult morbidity in the Whitehall II study" *Journal of Epidemiology and Community Health 55*(5), 301-307.

McCrary, J. and H. Royer (2011) "The Effect of Female Education on Fertility and Infant Health: Evidence from School Entry Policies Using Exact Date of Birth." *The American Economic Review* 101(1), 158-195.

Mörk, E., Svaleryd, H., and Sjögren, A. (2014). Hellre rik och frisk: Om familjebakgrund och barns hälsa. SNS förlag.

Nordlöf, B. (2001) "Svenska adoptioner i Stockholm 1918–1973 (Swedish adoptions in Stockholm 1918–1973)" FOU-rapport 2001:8 Socialtjänstförvaltningen, Stockholm Stad.

Oskarsson, S., K.-O. Lindgren and C. Dawes (2015) "It Runs in the Family". Mimeo. Department of Political Science, Uppsala University.

Osler, M., Nybo Andersen, A. M., Due, P., Lund, R., Trab Damsgaard, M., & Holstein, B. (2003) "Socio-economic position during childhood, birth weight, childhood cognitive function and mortality in adult life. A longitudinal study of Danish men born in 1953" *Journal of Epidemiology and Community Health*, *57*, 681-86.

Osler, M., Petersen, L., Prescott, E., Teasdale, T. W., and Sørensen, T. I. (2006). Genetic and environmental influences on the relation between parental social class and mortality. International Journal of Epidemiology, 35(5), 1272-1277.

Palme, M., and Sandgren, S. (2008). Parental income, lifetime income, and mortality. *Journal of the European Economic Association*, 6(4), 890-911.

Petersen, L., P. Kragh Andersen and T. Sørensen (2005) "Premature Death of Adult Adoptees: Analyses of a Case-Cohort Sample" *Genetic Epidemiology* 28, 376-382

Sacerdote, B. (2007) "How large are the effects from changes in family environment? A study of Korean American adoptees." *The Quarterly Journal of Economics* 122(1), 119-157.

Smith, K. P. and Christakis, N. A. (2008) "Social networks and health" *Annual Review of Sociology 34*, 405-429.

Smithers, L.G., Sawyer, A.C.P., Chittleborough, C.R.*et al.* (2018) "A systematic review and meta-analysis of effects of early life non-cognitive skills on academic, psychosocial, cognitive and health outcomes." *Nature Humamn Behav*iour 2, 867–880. https://doi.org/10.1038/s41562-018-0461-x

Solon, G. (1999) "Intergenerational Mobility in the Labor Market" in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics,* Volume III (Amsterdam: Elsevier) Chapter 29.

Sørensen, T., G. Nielsen, P. Kragh Andersen and T. Teasdale (1988) "Genetic and Environmental Influences on Premature Death in Adult Adoptees" *New England Journal of Medicine* 318(12), 727-732.

Socialstyrelsen (2009a) *The Swedish Cause of Death Registry.* Sweden, Stockholm. 2009.

Socialstyrelsen (2009b) *The Swedish National Patient Register. Inpatient diseases in Sweden 1987-2007.* Sweden, Stockholm. 2009.

Statistics Sweden (2012). "The Multi-generation register 2012. A description of contents and quality" Report BE96BR1301. Örebro: Statistics Sweden.

Appendix A: Additional results

Table A1. Associations between health indices (in percentile rank) and mother's and father's education (in years)

	(1) Non-adoptees	(2) Adoptees	(3) Adoptees	(4) Non-adoptees	(5) Adoptees	(6) Adoptees
	Hospitalization			Health index		
Education, Bio mother	0.361***	0.259*	0.523***	0.326***	0.222	0.437***
	(0.009)	(0.155)	(0.105)	(0.009)	(0.153)	(0.103)
Education, Bio father	0.499***	0.252*		0.453***	0.205	
	(0.007)	(0.130)		(0.007)	(0.128)	
Education, Ad mother		0.333**	0.248**		0.279**	0.238**
		(0.136)	(0.098)		(0.134)	(0.097)
Education, Ad father		0.176	0.371***		0.097	0.299***
		(0.111)	(0.078)		(0.110)	(0.078)
N	3,199,618	10,845	21,250	3,227,149	10,880	21,386

Note: Results from OLS regressions. Standard errors in parentheses; *** significant at 1%, ** at 5%, * at 10%. Each column represents a separate regression and all regressions include indicators for gender and birth cohort of children, and five-year intervals for parental cohorts. Column (1) and (4) contains all non-adopted children, and column (2) and (4) the sample of adoptees for which we can identify all parents, and column (3) and (6) adds adoptees with unknown biological fathers.

Table A2. Associations between hospitalization (in percentile rank) and parental education (in years), by child gender

	(1) Non-adoptees	(2) Adoptees	(3) Non-adoptees	(4) Adoptees
	Females		Males	
Education, Bio parents	0.843***	0.501*	0.922***	0.522**
	(0.011)	(0.261)	(0.010)	(0.249)
Education, Ad parents		0.449***		0.551***
		(0.173)		(0.167)
N	1,569,717	5,228	1,629,901	5,617

Note: Results from OLS regressions. Standard errors in parentheses; *** significant at 1%, ** at 5%, * at 10%. Each column represents a separate regression and all regressions include indicators for birth cohort of children, and five-year intervals for parental cohorts.

Table A3. Associations between health index (in percentile rank) and parental education, by child gender

	(1) Non-adoptees	(2) Adoptees	(3) Non-adoptees	(4) Adoptees
	Females		Males	
Education, Bio parents	0.705***	0.427*	0.892***	0.413*
	(0.011)	(0.256)	(0.010)	(0.246)
Education, Ad parents		0.214		0.502***
		(0.171)		(0.166)
N	1,578,905	5,239	1,648,244	5,641

Note: Results from OLS regressions. Standard errors in parentheses; *** significant at 1%, ** at 5%, * at 10%. Each column represents a separate regression and all regressions include indicators for birth cohort of children, and five-year intervals for parental cohorts.

Table A4. Parameter robustness with respect to inclusion of different sets of controls in the econometric model. Adoptees only. Dependent variable: Hospitalization.

	(1)	(2)	(3)	(4)	(5)	(6)
Education, Ad parents	0.5740***	0.4884***	0.3802***		0.4884***	0.3557***
	(0.1156)	(0.1197)	(0.1200)		(0.1197)	(0.1223)
Education, Bio parents		0.5066***	0.2975	0.6830***	0.5066***	0.4522**
		(0.1796)	(0.1818)	(0.1738)	(0.1796)	(0.1805)
Cohorts, Bio mother	No	Yes	Yes	Yes	Yes	Yes
Cohorts, Bio father	No	Yes	Yes	Yes	Yes	Yes
Cohorts, Ad mother	Yes	Yes	Yes	No	Yes	Yes
Cohort, Ad father	Yes	Yes	Yes	No	Yes	Yes
Health, Bio parents	No	No	Yes	No	No	No
Health, Ad parents	No	No	No	No	No	Yes
Region, Bio parents	No	No	Yes	No	No	No
Region, Ad parents	No	No	No	No	No	Yes
N	10,845	10,845	10,845	10,845	10,845	10,845
r2	0.0068	0.0104	0.0233	0.0071	0.0104	0.0166

Note: Results from OLS. Robust standard errors in parentheses; *** significant at 1%, ** at 5%, * at 10%. Each column represents a separate regression and all regressions include indicators for gender and birth cohort of children. Each column is adding parental characteristics.

Table A5. Associations between health indices (in percentile rank) and parental education (in years), using samples restricted to adoptees that move out from their municipality birth.

	(1) Hospitalization	(2) Health index
Education, Ad parents	0.543***	0.408***
	(0.129)	(0.127)
Education, Bio parents	0.591***	0.483**
	(0.194)	(0.190)
N	9,304	9,333

Note: Results from OLS regressions. Standard errors in parentheses; *** significant at 1%, ** at 5%, * at 10%. Each column represents a separate regression and all regressions include indicators for gender and birth cohort of children, and five-year intervals for parental cohorts. The sample only include adoptees who was adopted by a mother who lived in a different municipality in 1960 than the biological mother of the adoptee.

Table A6. Associations between health indices (in percentile rank) and parental education (in years), using samples restricted to adoptees that adopted out at least two children. Specifications w/o and with bio mother fixed effects.

	(1) Hospitalization	(2) Hospitalization FE	(3) Health index	(4) Health index FE
Education, Ad mother	0.845***	0.838	0.939***	1.206
	(0.274)	(0.763)	(0.268)	(0.808)
Education, Bio mother	0.292	0.000	0.283	0.000
	(0.378)	(.)	(0.364)	(.)
N	2,254	2,254	2,261	2,261

Notes: Robust standard errors in parentheses. All models include controls for gender and birth cohorts.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A7. Interaction effects between measures of educational attainments (in years) for biological and adopting parents

	(1) Hospitalization	(2) Health index
Bio*Ad parents, education	0.015	0.009
	(0.059)	(0.059)
Education, Bio parents	0.361	0.331
	(0.610)	(0.603)
Education, Ad parents	0.358	0.272
	(0.531)	(0.531)
N	10,845	10,880

Note: Results from OLS regressions. Standard errors in parentheses; *** significant at 1%, ** at 5%, * at 10%. Each column represents a separate regression and all regressions include indicators for gender and birth cohort of children, and five-year intervals for parental cohorts.

Table A8. Association between mortality and mother's and father's education (in years) (hazard ratios)

	(1) Non-adoptees	(2) Adoptees	(3) Adoptees
Education, Bio mother	0.978***	1.028	0.993
	(0.001)	(0.019)	(0.012)
Education, Bio father	0.971***	1.006	
	(0.001)	(0.015)	
Education, Ad mother		0.967*	0.973**
		(0.017)	(0.012)
Education, Ad father		0.989	0.996
		(0.013)	(0.009)
N	3,227,149	10,880	21,386

Note: Results from Cox proportional hazard models. Standard errors in parentheses; *** significant at 1%, ** at 5%, * at 10%. Each column represents a separate regression and all regressions include indicators for gender and birth cohort of children, and five-year intervals for parental cohorts. Column (1) contain all non-adopted children, column (2) the sample of adoptees for which we can identify all parents, and column (3) adds adoptees with unknown biological fathers.

Table A9. Association between mortality by cause of death and parental education (in years), (hazard ratios)

	(1) Cancer	(2) Circulatory	(3) Preventable	(4) Treatable
Non-adoptees				
Education, Bio parents	0.9617***	0.9601***	0.9312***	0.9218***
	(0.0016)	(0.0023)	(0.0041)	(0.0035)
Mean dep. Var	0.044	0.015	0.007	0.011
Observations	3,227,149	3,227,149	3,227,149	3,227,149
Adoptees				
Education, Bio parents	1.0471	1.0622	1.0401	1.0371
	(0.0400)	(0.0485)	(0.1083)	(0.0637)
Education, Ad parents	0.9235***	0.9836	0.8766*	0.9703
	(0.0280)	(0.0279)	(0.0647)	(0.0552)
Mean dep. Var	0.025	0.018	0.005	0.007
N	10,880	10,880	10,880	10,880

Note: Results from Cox proportional hazard models with competing risks. Standard errors in parentheses; *** significant at 1%, ** at 5%, * at 10%. Each column represents a separate regression and all regressions include indicators for gender and birth cohort of children, and five-year intervals for parental cohorts. Panel A is based on a sample of non-adopted children, and Panel B on the sample adoptees for whom we have information on all parents. Diagnoses codes ICD10: Cancer C00-D48, Circulatory I00-I99, Preventable C33-C34, K70, K74.3-K74.6, and Treatable A15-A19, B90, C53, I05-I09, J00-J99, J45, J46, K35-K38, K40-K46, I10-I15, I60-I69, K80-K81.

Table A10. Diagnoses codes for different diagnose categories

Diagnoses	ICD10 codes
Cancer	C00-D48
Circulatory	I00-I99
Preventable	C33-C34, K70, K74.3-K74.6
Treatable	A15-A19, B90, C53 , I05-I09, J00-J99, J45, J46, K35-K38, K40-K46, I10-I15, I60-I69, K80-K81

Table A11. Mediation analysis: Health index

	(1) Health index	(2) Health index	(3) Health index	(4) Health index	(5) Health index	(6) BMI	(7) Physical fitness	(8) Education	(9) Non-cognitive	(10) Cognitive
Education, Ad parents	0.309	0.224	0.065	0.000	-0.126	-0.058**	0.020***	0.138***	0.049***	0.071***
	(0.190)	(0.190)	(0.191)	(0.191)	(0.191)	(0.024)	(0.007)	(0.014)	(0.007)	(0.006)
Education, Bio parents	0.816***	0.728**	0.516*	0.388	0.278	-0.023	0.037***	0.169***	0.064***	0.102***
	(0.285)	(0.284)	(0.284)	(0.287)	(0.285)	(0.031)	(0.010)	(0.022)	(0.010)	(0.010)
BMI		-0.791***			-0.755***					
		(0.171)			(0.165)					
Physical fitness		1.913***			0.240					
		(0.479)			(0.513)					
Education			1.769***		1.018***					
			(0.228)		(0.253)					
Non-cognitive				2.686***	2.507***					
				(0.480)	(0.513)					
Cognitive				2.515***	1.755***					
				(0.511)	(0.540)					
N	3,944	3,944	3,944	3,944	3,944	3,944	3,944	3,944	3,944	3,944

Note: Results from OLS regressions. Standard errors in parentheses; *** significant at 1%, ** at 5%, * at 10%. Each column represents a separate regression and all regressions include indicators for gender and birth cohort of children, and five-year intervals for parental cohorts. The sample is based on male adoptees with information on educational attainment and measures on health and skills from military conscription.

Table A12. Mediation analysis: Health index. Both female and male adoptees.

	(1) Health index	(2) Health index	(3) Education
Education, Ad parents	0.363***	0.166	0.152***
	(0.119)	(0.120)	(0.009)
Education, Bio parents	0.398**	0.160	0.183***
	(0.177)	(0.178)	(0.014)
Education		1.299***	
		(0.131)	
N	10,812	10,812	10,812

Note: Results from OLS regressions. Standard errors in parentheses; *** significant at 1%, ** at 5%, * at 10%. Each column represents a separate regression and all regressions include indicators for gender and birth cohort of children, and five-year intervals for parental cohorts. The sample is based on all adoptees with information on educational attainment.