Deutsches Institut für Wirtschaftsforschung

DIW BERLIN

www.diw.de

SOEP The German Socio-Economic Panel Study

SOEPpapers
on Multidisciplinary Panel Data Research

375

Daniel Kemptner • Jan Marcus

Spillover Effects of Maternal Education on Child's Health and Schooling

Berlin, April 2011

ISSN: 1864-6689 (online)

German Socio-Economic Panel Study (SOEP)
DIW Berlin
Mohrenstrasse 58
10117 Berlin, Germany

Contact: Uta Rahmann | soeppapers@diw.de

Spillover Effects of Maternal Education on Child's Health and Schooling

Daniel Kemptner[§], and Jan Marcus[#]

April 20, 2011

Abstract

This is the first study investigating the causal effect of maternal education on child's health and schooling outcomes in Germany. We apply an instrumental variables approach that has not yet been used in the intergenerational context. For that purpose, we draw on a rich German panel data set (SOEP) containing information about three generations. This allows instrumenting maternal education by the number of her siblings while conditioning on a set of variables describing the grandparents' social status and the area where the mother grew up. Given these variables, the number of siblings generates exogenous variation in the years of education by affecting the household resources available per child. We present evidence for strong and significant effects on schooling outcomes for both sexes. And, we find substantial effects on health behaviour for adolescent daughters, but not for adolescent sons. We show that possible concerns for the validity of the instrument are unlikely to compromise these results. We also discuss assortative mating and household income as possible channels of causality.

Keywords: Intergenerational mobility, returns to education, health, instrumental variables

JEL-Codes: C26, I12, I21, J62

Valuable comments by Irma Clots-Figueras, Rafael Lalive, Adam Lederer, Frauke Peter, Nils Saniter and Thomas Siedler as well as by participants at SOEP Brown Bag, FU Berlin, the Berlin Leibniz Seminar on Labor Research the 7th International Young Scholar German SOEP Symposium in Delmenhorst and the Spring Meeting of Young Economists in Groningen are gratefully acknowledged.

[§] DIW Berlin, Graduate Center, Mohrenstr. 58, 10117 Berlin, Germany. Email: dkemptner@diw.de
[#] DIW Berlin, Graduate Center, Mohrenstr. 58, 10117 Berlin, Germany. Email: jmarcus@diw.de

1. Introduction

When analyzing returns to education, economists often focus on wages and income (see Card, 1999, for an overview). In recent years, research concentrated also on the causal effect of education on non-market outcomes like health (see Cutler/Lleras-Muney, 2008, and Grossmann, 2006, for overviews). Furthermore, researchers point to the intergenerational spillover effects of education (Black/Devereux, 2010, and Currie, 2009, provide overviews), although, most research on the intergenerational effects of education concentrates on correlations. Only a few studies analyze the causal effects of parental education on child's health and schooling outcomes (see Black/Devereux 2010). Quantifying such intergenerational links is not only relevant regarding optimal investments into education, but also relates to social mobility. The more a child's outcomes are determined by its parents' education, the less socially mobile a society can be considered.

Our paper is the first study investigating the causal effect of maternal education on child's health behaviour and schooling outcomes in Germany. Thus, we study the intergenerational transmission of human capital by looking at the effects on measures for the production of education and health of the child. While schooling success indicates formation of education, health behaviour relies to health production (see Grossman 2006). Education and health constitute main components of human capital. We look at five different binary outcome variables: current smoking status, overweight, frequency of doing sports, grade repetition and attending the highest secondary schooling track. We apply an instrumental variables approach that has not yet been used in the intergenerational context. For this purpose, we draw on a rich German panel data set, the Socio-Economic Panel Study (SOEP), containing information about three generations. This allows instrumenting maternal education by the number of her siblings, while conditioning on a set of variables describing the grandparents' social status and the size of the area where the mother grew up. Given these variables, the number of siblings generates exogenous variation in the years of education by affecting the household resources available for educational investments per child. This assumes implicitly that parents are constrained in borrowing against their children's future earnings. The first stage of our instrumental variables approach can be interpreted as a test for this assumption.

Studies on the causal effects of parental education on child's health and schooling outcomes in developed countries produce mixed evidence. Using instrumental variable (IV) approaches, Oreopoulos, Page and Stevens (2003, 2006), Maurin and McNally (2008), as well as Carneiro, Meghir and Parey (2007) find negative causal effects of parental education on grade repetition, while Björklund, Lindahl and Plug (2006) and Holmlund, Lindahl and Plug (2008) obtain insignificant effects on child's years of schooling. For newborns, Currie and Moretti (2003) find maternal education to reduce the risks of low birth weight and preterm birth in a causal way. This finding is not corroborated by the IV-study of McCrary and Royer (2011). For teenage children, Carneiro et al. (2007) and Lindeboom, Llena-Nozal and van der Klaauw (2009) find no significant effects of parental education on the children's health status in their IV-analyses.

Previous studies for Germany analyze the intergenerational correlation of education (Heineck/Riphahn 2007) and of health (Coneus/Spieß 2008), as well as the correlation between parental education and child health (Lamerz et al. 2005), but not the causal effects. We add to the literature by analyzing the causal effects of education on various measures of health-related behaviour and schooling outcomes. We focus only on mothers because the SOEP reports on the partner of the mother and not on the biological father.

We find strong and significant causal effects of maternal education on schooling outcomes for both sexes as well as on health behaviour for adolescent daughters. A unit increase in maternal years of education is estimated to reduce the child's probability to repeat a grade by 7.4 percentage points, and increases the child's probability of pursuing the highest secondary schooling track by 9.5 percentage points. The findings on schooling outcomes do not differ by the sex of the child. Furthermore, we estimate that an increase in maternal education by one year reduces the adolescent daughter's probability to smoke at age 18/19 by 5.8 percentage points and increases the adolescent daughter's likelihood of doing sports at least once a week by 6.7 percentage points. However, we do not obtain significant effects of maternal education on overweight for both daughters and sons. There is no effect on sons' health behaviour in general. These findings are robust over a variety of specifications. We also discuss assortative mating and household income as possible channels of causality. Our findings suggest that assortative mating is relevant. It seems that the mother's partner is more relevant for sons than for daughters. Together with the fact that we only find effects on the health behaviour of daughters, this suggests that mothers are more influential with respect to daughters. Household income does not explain the effects.

The paper is structured as follows. Section 2 describes the data and presents descriptive statistics. Section 3 contains a detailed discussion of our empirical strategy. In section 4, we present both Probit and IV- Probit results and we compare our results with other findings from the literature on intergenerational transfers. Section 5 presents various sensitivity checks. Subsection 5.1 deals with possible concerns for the validity of our instrument and presents evidence that these concerns are unlikely to compromise the results. Subsection 5.2 checks the robustness of our findings with respect to model specification. Section 6 investigates channels of causality like assortative mating and family income. Section 7 concludes with a discussion on the implications of our findings.

2. Data and Descriptive Statistics

2.1 The Sample

In our analysis we make use of the rich data from the German Socio-Economic Panel Study (SOEP). The SOEP started in 1984 and annually collects information at the household and individual levels (see Wagner, Frick and Schupp 2007). In 2009 more than 12,000 households participated in this panel study.

The SOEP is the best data set for our purposes: Not only is it one of the largest and longest-running panel studies in the world, it also provides detailed information on health behaviour and schooling success of adolescents. Furthermore, due to the collection of additional biographical information of adult respondents, for these adolescents data on their parents *and* on their grandparents are available.

The sample consists of West German children born between 1983 and 1991, using data from when these children were around 18 years of age. Hence, the sample is pooled across survey years. For the schooling outcomes we draw on data from a special youth questionnaire that the adolescents answered in the year they turned 17. For health related variables, we use data from the year when the adolescents answered the relevant questions on the individual adult questionnaire for the first time. Since some of the health variables are only included in even numbered years, for some adolescents we use information from the year they turned 18 and for the rest we use information from the year they turned 19. In the regressions we control for these age differences through fixed birth year effects.

2.2 Outcome Variables

We construct a variable “overweight” indicating a body mass index (BMI) greater than 25. We code a binary variable “currently smoking” according to the question “Do you currently smoke, be it cigarettes, a pipe or cigars?” The SOEP started asking health questions, including weight and smoking behaviour, in even numbered years, starting with 2002.[1] A variable on sport activities indicates whether an adolescent is doing sports at least once a week.[2] We

[1] The SOEP collects data on smoking behaviour also in the years 1998, 1999 and 2001. These questions, though, differ in their phrasing. Therefore, we exclude information from these survey years.
[2] We computed the regressions also for a slightly different definition of this variable. The results differ only marginally, when we consider a person as being active who is doing sports at least once a *month*. These and other results not shown are available from the authors upon request.

generally use information on sport behaviour from the year they turned 18. However, this variable was not collected in the survey years 2002, 2004 and 2006. For those who turn 18 in these years, we use the information about doing sports from the year they turned 19 – information gathered during the next wave.

In addition to the measures of health behaviour, we look at two variables that measure success in school. Both variables build on questions in a special questionnaire SOEP uses when adolescents turn 17. Because this questionnaire only started in 2000, we can only include adolescents born in 1983 and on in the analyses. The first of the school outcome variables is an indicator of whether the child ever repeated a grade. It is directly asked from the adolescents and *not* constructed from actual/normal grade for age comparisons. Grade repetition is not only an indicator for individual school failure but it also has high financial costs for society (Jackson 1975).

After primary school, the German school system selects children into one of three tracks: basic track (*Hauptschule*), intermediate track (*Realschule*), or academic track (*Gymnasium*). Pupils can only obtain the *Abitur* from academic track schools. The *Abitur* is the diploma usually required for matriculating into a German university. We construct a variable indicating whether an adolescent is in an academic track school at age 17. Since students cannot graduate from an academic track school before turning 18, this schooling measure captures all children who will presumably acquire the Abitur.

2.3 Parental and Grandparental Variables

At the parental level, we focus only on mothers because the SOEP collects data on the mother's partner, who may or may not be the biological father of the child. Relevant data for mothers include years of education, number of siblings and population of the area the mother grew up until the age of 15. The SOEP constructs the years of education variable from the respondents' information about the obtained level of education and adds time for additional occupational training.[3] For the numbers of siblings, we use the earliest available information about brothers and sisters collected in the survey.[4] Since siblings might have died, this is the best approximation of the number of brothers and sisters when the mother went to school. The area where the mother grew up is a discrete variable with four categories according to the size of the hometown: countryside, small city, medium city and large city. To control for possible time trends in the number of siblings and the years of education we also include the mothers' age at birth of the child.[5] All information about mothers is self-reported by the mothers.

At the level of the parents of the mother, we use data on the educational level. For both, grandfathers and grandmothers, we construct dummy variables according to five educational levels: secondary school degree, intermediate/technical school degree, general university-entrance diploma, other degree and no school degree/no school attended. Information on the grandparents is either contributed by the grandparents directly (less than 5 percent) – if they are SOEP participants – or by proxy via interviews of the mothers: All individuals with a valid personal interview in the SOEP are requested to answer the supplementary biography questionnaire with questions on their parents and their social origin. Missing values at the grandparental level are imputed as described in the appendix.

2.4 Descriptive Statistics

Table 1 displays unweighted means and standard deviations for relevant variables at both the maternal and child level. The sample consists of West German children born between 1983 and 1991 and excludes children whose mothers were educated in the German Democratic

[3] If the variable years of education is missing for an individual in a given survey year, we use information from other survey years.
Following Kemptner et al. (2011), we employed also a different measure of the years of education, in which we only considered years of primary and secondary schooling: 9 years for individuals without school degree and those with basic track degree, 10 years for those with intermediate track school or other degree, 12 years for technical school degree and 13 years for general university-entrance diploma. However, the Probit and IV-Probit results did not change qualitatively, only the size of the coefficient estimates increased.

[4] The SOEP collects this information in 1990, 1996, 2001, 2003 and 2006. We consider siblings inside and outside of the household.

[5] More specifically, we put five subsequent birth years in one category and include binary variables for these birth year categories (before 1950, 1950-1954, 1955-1959, 1960-1964, 1965 and after).

Republic. The last two columns exclude mothers *without* siblings. This additional restriction is imposed to allow for a good linear approximation between the number of siblings and maternal education, as explained in the next section in more detail. Table 1 shows that excluding mothers without siblings does not alter the marginal distribution of the outcome variables. About 29 % of adolescent children smoke, more than 17 % are overweight, 55 % are doing sports at least once a week, 39 % are pursuing an academic track and 24 % ever repeated a grade. Excluding mothers without siblings reduces the sample size by 14 %.

Figure 1 presents graphs of the local polynomial regression of order zero, i.e. mean smoothing, between mother's years of education and various child outcomes. For almost all outcome measures there is a monotonous relationship: Positive health behaviour and educational success of the child increase almost linearly with the mother's education. For instance, the chance of doing sports at least once a week is around 40 % for children of poorly educated mothers, 50 % for children of mothers with about 10 years of education and 70 % for children of mothers with more than 15 years of education. We find steep maternal education gradients for the child's educational variables. The chances of being in an academic track school at age 17 are eight times higher for children of the best educated mothers in comparison to children of the least educated mothers. The increase in the overweight probability on the right tale of the education distribution is not statistically significant as the confidence bands indicate. The next section describes our empirical strategy to determine whether these bivariate relationships constitute indeed causal relationships.

3. Empirical strategy

We estimate the effect of maternal years of education on binary child's outcomes. For this purpose, we estimate single (Probit) and two-equation models (IV-Probit).[6] All models are estimated with robust standard errors that are clustered by mothers, accounting for serial

[6] In the section on robustness checks, we also present results from a 2SLS-model. Being more robust regarding the distributional assumptions of the error terms, the estimated effects differ only marginally.

correlation between children of the same mother. Our single equation model linking child's outcome to maternal years of education is specified as follows:

$$H_c = 1\left[\beta_0 + \beta_1 \times S_m + \pi_{cohort,m} + \eta_{origin,m} + \delta_{educ,gp} + \nu_{cohort,c} + \lambda_{sex,c} + \varepsilon_c > 0\right] \quad \textbf{(1)}$$

where H_c is child's outcome and S_m is maternal years of education. $\pi_{cohort,m}$, $\eta_{origin,m}$, $\delta_{educ,gp}$, $\lambda_{sex,c}$, and $\nu_{cohort,c}$ are sets of fixed effects accounting for mother's birth cohort, the size of the area where the mother grew up, grandparents' level of education, child's birth cohort, and child's sex. ε_c is an idiosyncratic child specific error term that is normally and identically distributed. $1[\cdot]$ is an indicator function.

Estimating **(1)** as a single equation model will only produce consistent parameter estimates if maternal years of education, S_m, are uncorrelated with ε_c. Since maternal years of education are likely to be correlated with unobserved confounders, we expect the coefficient estimates to be biased in an unknown direction.

The endogeneity of S_m can be dealt with by instrumenting S_m with Z_m, where Z_m must meet the following two conditions:

$$E\left(\varepsilon_c \mid Z_m\right) = 0 \quad \textbf{(validity)}$$

$$E\left(S_m \mid Z_m, X\right) \neq E\left(S_m \mid X\right) \quad \textbf{(relevance)}$$

where X contains all other covariates. Z_m is a valid instrument if it affects the child's outcome only through mother's years of education, given the other covariates. Z_m is a relevant instrument if the explanatory power of Z_m with respect to S_m is sufficiently large, given the other covariates. Various instruments for education are proposed in the literature on returns to education (see Card, 1999, and Grossman, 2006, for overviews). A first wave of IV studies relies on family characteristics as instruments, such as parents' income and parents' schooling. While these instruments are strongly associated with education, the validity assumption seems questionable. A second wave of IV studies use variations in educational policies and other natural experiments. This second wave of IV studies faces less criticism regarding the validity assumption. However, the association with education is often weak and, hence, weak instrument problems may arise. Researchers frequently draw on huge sample sizes to mitigate this problem. A drawback of huge data sets is that these often do not include detailed outcome measures.

Another problem with policy changes and other natural experiments is that they only affect certain cohorts. For Germany some natural experiments are used to study the returns to education. Reinhold and Jürges (2010) rely on changes in the abolition of secondary school fees; Pischke and von Wachter (2008) as well as Kemptner, Jürges and Reinhold (2011) make use of changes in compulsory schooling that occurred at different times in the German states; Jürges, Reinhold and Salm (2009) rely on school openings during the educational expansion in Germany. These studies use data from the German Microcensus, which, however, lacks detailed information on child outcomes. These natural experiments do not affect our cohorts of mothers and/or have week first stage properties when used to instrument the years of education for mothers in the SOEP.

We do not rely on policy changes but instead use the number of mother's siblings as an instrument for maternal education. This identification strategy works also for cohorts unaffected by policy changes and for the limited sample sizes of common household panels. This instrument was suggested before (e.g. Sander 1995) and it may look like a typical instrument from the first wave, thus suffering from validity problems. We improve this approach in various ways. First, we keep other characteristics of the mother's family constant. This rationale is also used by Gebel and Pfeiffer (2010) when estimating monetary returns to education for wage earners in Germany. Fertility is higher on the countryside and negatively correlated with social status. Therefore, we control for the social status of the grandparents and the size of the area where the mother grew up. The grandparent's social status is controlled for through a set of education fixed effects accounting for heterogeneity between the levels of education of grandmother and grandfather. The area where the mother grew up is accounted for by the inclusion of dichotomous variables according to the four categories previously described. Second, the validity assumption does not require that the mother's outcomes are unaffected by the number of her siblings. Only the *child's* outcomes must be uncorrelated with the number of mother's siblings. Hence, possible concerns about the validity of the instrument always need to involve some kind of inheritance. Third, in section 5 we show that various concerns about the validity of this instrument do not compromise our results.

The number of the mother's siblings should also be a relevant instrument because the amount of resources that is available per child for investments into education depends substantially on the number of children in the household. This assumes that parents are constrained in borrowing against their children's future earnings. A significant effect of the number of mother's siblings on her education in the first stage points to such a borrowing constraint of the grandparents. Even though there are no schooling fees and very low or no tuition fees at public educational institutions in Germany, investments into children's education involve forgone earnings for both the parents and the children. Parents' time constraints and limited housing space may impose pressure upon the children to make their own living instead of spending more time on educational investments. Table 2 contains a regression of maternal years of education on a set of dummy variables indicating the number of siblings. We estimate two specifications: with and without demographic controls.[7] The table shows a clear negative relationship between maternal education and the number of siblings. However, mothers without siblings (omitted category) seem to be special having on average less education than mothers with one or two siblings. Black, Devereux and Salvanes (2005) also find this only child particularity, which disappears when they consider the intact family subsample. For this reason we exclude children of mothers without siblings from our sample to allow for a good linear approximation of the relationship between the number of

[7] These demographic variables include controls for child's sex, child's year of birth, mother's age at birth of the child, fixed effects for the area the mother grew up and for the educational levels of the mother's parents.

siblings and maternal education. In subsection 5.2 we present robustness checks for the inclusion of children of mothers without siblings and for other specifications of the functional relationship in the first stage (log-specification, dichotomous variables indicating the number of siblings).

We implement the IV strategy by estimating the following two-equation model by the method of maximum likelihood:

$$S_m = \gamma_0 + \gamma_1 \times Z_m + \omega_{cohort,m} + \kappa_{origin,m} + \rho_{educ,gp} + \tau_{cohort,c} + \xi_{sex,c} + u_m \qquad \textbf{(2)}$$

$$H_c = 1\left[\beta_0 + \beta_1 \times S_m + \pi_{cohort,m} + \eta_{origin,m} + \delta_{educ,gp} + \nu_{cohort,c} + \lambda_{sex,c} + \varepsilon_c > 0\right] \qquad \textbf{(3)}$$

u_m and ε_c are idiosyncratic error terms being bivariate normally and identically distributed. Under the assumptions of instrument validity and relevance, joint estimation of (2) and (3) as an IV-Probit model produces consistent parameter estimates. The coefficients of the first stage can be directly interpreted as marginal effects. The parameters of a Probit model cannot be given the interpretation of marginal effects. For this reason, we compute average marginal effects and apply the delta method when calculating standard errors.

4. Results

Table 3 contains the first stage results of the IV-Probit model. Table 4 presents the second stage results of the IV-Probit model and the findings of the single equation Probit model that serves as a benchmark when interpreting the IV estimates. The IV-Probit model is estimated with two specifications. The first specification does not take into account the characteristics of grandparents' household (level of education, size of the area where the mother grew up) at the time when the mother was a child. As has been argued above, the validity of our instrument is much more credible when these characteristics are controlled for. The second specification is our baseline specification that is outlined in the previous section. The single equation Probit model is estimated for specification 2.

4.1 Probit results

In the Probit models, the average marginal effects for indicate a significant association between maternal education and children's outcomes. One more year of mother's education is associated with a decrease in the child's probability of being a smoker at age 18/19 by about 1.7 percentage points. The observed relationship seems to be stronger for daughters than for sons. The likelihood of being overweight at age 18/19 is associated with a decrease by about 1 percentage point per year of maternal education. This association is not significantly different for daughters and sons. Furthermore, there seems to be a strong relationship between maternal education and the child's likelihood of doing sports at least once per week. The estimates suggest that each additional year of maternal education is associated with an increase in the probability of doing sports regularly by 3.4 percentage points.

Turning to the relationship between maternal education and the child's schooling outcomes, the intergenerational association seems to be even larger. Each additional year of maternal education is associated with an increase in the probability of being on an academic track school by 6.8 percentage points. Furthermore, there is a significant association between maternal education and the probability that a child must repeat a year. The sex of the child seems not to matter for the estimated educational relationships.

4.2 IV-Probit results

4.2.1 First stage

Table 3 presents the first stage coefficients of the IV estimation for the specification without grandparental control variables (specification 1) and for the baseline specification (specification 2), which includes the grandparents' levels of education and size of the area where the mother grew up as controls. The estimated effects of the number of mother's siblings on her educational attainment are highly significant. An additional sibling decreases the years of education by 0.2-0.3 years for the baseline specification and between 0.3-0.4

years for specification 1. This shows that conditioning on characteristics of the grandparents' household reduces, *ceteris paribus*, the relationship between the number of siblings and educational achievement of the mothers. The small differences in the first stage coefficient estimates according stem from different sample sizes for the outcome measures. Regarding significance of the relationship, we are mainly interested in the respective F-statistics for specification 2. Testing the assumption for specification 2 that there is no effect of the instrument on maternal education, given the other covariates, all our F-statistics exceed 42 for the pooled sample. All the F-statistics are above 12 for the sample of sons, and above 33 for the sample of daughters. Thus, there is no concern about a weak instruments problem. The size of the F-statistics depends heavily on the sample size. The estimated significant effects in the first stage point to financial constraints of the grandparents when investing into their daughter's education.

4.2.2 Second Stage

We find large and significant causal effects of maternal years of education on daughter's smoking and sport behaviour in both IV specifications. However, we do not find any significant causal effects on son's health outcomes in general. There is also no significant effect on overweight for girls. In our baseline specification, specification 2, the probability of the daughter doing sports regularly is increased by 6.7 percentage points per year of maternal education. In addition, one additional year of maternal education decreases the likelihood of the daughter being a smoker at age 18/19 by 5.8 percentage points. Also Loureiro, Sanz-de-Galdeano and Vuri (2010) find that mothers are only influential with respect to the smoking behaviour of daughters but not of sons. These gender differences are in line with the idea of gender-specific parental role-models and the finding that children identify stronger with the same-sex parent (Starrels 1996).

It is interesting to see that the estimates are very similar and mostly smaller in specification 1. This does not only underline the need to control for the grandparents' household characteristics, it also suggests that failure to do so leads to an underestimation of the true effects.

While the effects on child's health outcomes differ depending on child's sex, we find large, significant, and very similar causal effects of maternal education on schooling outcomes for both sons and daughters. Our estimates suggest that one additional year of maternal education increases the likelihood of her child being in an academic track school at age 17 by 9.5 percentage points. Further, one additional year of maternal education decreases the probability that a child must repeat a year by 7.4 percentage points.

The fact that we find very similar effects for sons and daughters on schooling outcomes suggests that the child's education does not explain the effects on daughters' health behavior. The very large causal effects on the children's schooling outcomes indicate that the German education system involves a high degree of intergenerational persistence of education. Since we control for unobserved heterogeneity, the size of these effects is all the more alarming. Comparing the effects of maternal education on child's schooling outcomes with the estimates from the Probit models, the estimated effects from the IV-Probit models are, in general, much larger. This is in line with previous findings on intergenerational education transmission. Oreopoulos et al. (2003, 2006), Carneiro et al. (2007) and Maurin and McNally (2008) also find larger effects when instrumenting parental education. Three factors might be responsible for this finding. First, measurement error in maternal education attenuates the Probit estimates. Second, unobserved variables that are negatively correlated with maternal education but positively with better child outcomes might result in downward biased estimates. Third, in the presence of effect heterogeneity, IV approaches do not identify the average effect for the overall population but rather local average effects for the so-called compliers. In our case, these compliers are those mothers whose educational attainment was affected by their number of siblings.

For overweight also Lindeboom et al. (2008) do not find a significant causal effect for various ages of the children. The magnitude of our estimated effects on grade repetition is in the range of previous findings. Carneiro et al. (2007) use exogenous variation in education induced by variation in schooling costs in the US. For 7-8 year old white children they estimate a decrease in the probability to repeat a great of about 3 percentage points for each additional year of maternal education. The effects for children aged 12-14 are very similar, but stronger for girls than boys. For 12-14 year old black children, Carneiro et al, estimate an even larger causal effect: 6.4 percentage point reduction for each additional year of maternal education. Maurin and McNally (2008) only identify the causal effect of paternal education. For France they find a reduction in the probability of grade repetition by 33 percentage points. Oreopoulos et al. (2003, 2006) exploit changes in compulsory schooling in the US. In the 2006 paper they only report effects of joint parental education. However, in the 2003 paper, they present estimates for maternal education separately. They estimate that one additional year of maternal education reduces the probability of grade repetition by 5 percentage points for the overall population of 15-16 year olds and by more than 6 percentage points for 15-16 year old children of mothers with less education.

Overall, our estimates point to strong intergenerational transfers of human capital in Germany. While the effects on child's health outcomes differ with the sex of the child, we find large, significant, and very similar causal effects of maternal education on the schooling outcomes of both sons and daughters.

5. Sensitivity checks

The consistency of our estimates rests on the assumption that the instrument identifies exogenous variation in the endogenous education variable, given the other covariates. There are four arguments that can be brought forward against the validity assumption. First, controlling for the grandparents' level of education and the area where the mother grew up may not sufficiently account for the social status of the grandparents. This could violate the validity assumption if social status is correlated with both fertility and the grandchild's outcomes. Second, the financial constraints induced by the number of siblings could also constrain investments into the mother's health, which might have direct effects on the child. Third, the financial constraints induced by the number of siblings could make the grandparents move to a bad area where negative peer group effects might still affect the children's outcomes. Fourth, the grandparents' fertility could affect mother's fertility and lead to financial constraints having direct impacts on the child's outcomes.

In subsection 5.1, we provide evidence that all four concerns are unlikely to compromise the validity of our approach. Table 5 displays the associated findings. We only consider outcome variables with significant effects in the baseline specification. Furthermore, subsection 5.2 checks the robustness of our findings with respect to model specification. There, we consider both functional form assumptions and distributional assumptions regarding the error terms of the IV-Probit model.

5.1 Validity assumption

a) Grandparents' social status may not sufficiently be controlled for

We re-estimate the baseline IV model and include the logarithm of the International Socio-Economic Index of Occupational Status (ISEI) and a variable indicating migration background (direct or indirect) as additional controls for grandparents' social status (specification 3). The ISEI assigns scores to almost 300 different occupation categories "in such a way as to maximize the role of occupation as an intervening variable between education and income" (Ganzeboom, de Graaf and Treiman 1992: 2). Combined with the educational level of the grandparents, the ISEI score is a way to control for the income of the grandparents. The ISEI score is derived from the occupational status of grandfather and grandmother. SOEP questions on the occupational status of grandfather and grandmother are formulated to reflect the situation when the mother was 16. For each pair of grandparents we make use of the highest ISEI score, which is in most cases the score of the grandfather. Missing values are imputed as described in the appendix.

Including the ISEI score and a migration dummy only leads to marginal changes in the estimated effects and to a slight increase in the size of the standard errors. It is probably due to this loss in precision that the effect on the daughter doing sports regularly becomes insignificant. Since additional controls for the grandparents' social status do not change the results, we are confident that grandparents' social status is sufficiently controlled for in our baseline specification.

b) Siblings affect investments into mother's health

If the number of siblings affects mother's health and if mother's health directly affects child outcomes, the validity assumption of our instrument would be violated. We deal with this concern by including additional controls for mother's health into our baseline model (specification 4). As controls for maternal health, we use BMI, a variable indicating the mother being a smoker, as well as two summary measures for physical and mental health. The latter two variables, the physical and mental summary scale, are two indices that both constitute a combination of several survey questions that are weighted according to a specific algorithm (Andersen et al. 2007). When estimating the IV model, we only find marginal changes in the estimated effects.

c) Siblings make grandparents move to a bad area

Long-lasting negative peer group effects could be induced by the number of siblings through financial constraints and the decision to move into a bad neighbourhood. We deem that this should be primarily a problem in large cities where neighbourhoods are very heterogeneous and where the quality of a child's elementary school depends largely on the neighbourhood. For this reason, we re-estimate the baseline specification excluding mothers who grew up in large cities (specification 5). Most estimated effects change only slightly. However the effect of maternal education on the probability that a child must repeat a grade becomes smaller and insignificant for daughters. Furthermore, the effect on sport behaviour for daughters also becomes insignificant although the corresponding effect for the pooled sample remains significant. The effects are estimated less precisely because of the smaller sample size. Slight differences between the estimated effects for the whole sample and the subsample may also result from effect heterogeneity between large cities and less urbanized areas.

d) Intergenerational correlation of fertility

Grandparents' fertility could affect mother's fertility and lead to financial constraints that have a direct impact on child outcomes. Indeed, we find that the number of mother's siblings explains 4% of the variation in the number of mother's own children. We address this concern by including fixed effects for the number of mother's children as additional controls in our IV model (specification 6).[8] When estimating the model, we only find marginal changes in the estimated effects. Hence, all four concerns for the validity of our instrument are unlikely to compromise our results.[9]

[8] More specifically, we include dummy variables for 1, 2, 3 and 4 or more children.

[9] As a further means to test the instrument validity, we estimate the causal effect of mother's education on her own health behaviour and compare the results to previous research on Germany using policy changes. Jürges et al. (2009) instrument education with the density of academic track schools in each state. Kemptner et al. (2011) make use of the fact that a change in compulsory schooling occurred at different times in each German state. Both studies use the German Microcensus. We focus the comparison on the variables currently smoking, overweight (BMI>25) and obese (BMI>30) since these are the only variables that are included on both the SOEP and the Microcensus. Making the sample as comparable as possible in terms of birth cohorts, we obtain that all effects that are found to be significant in these studies are also significant in our replication: Jürges et al. (2009) find a significant reduction of 6.1 percentage points in the smoking probability for each year of maternal education, while we estimate a reduction of 6.7 percentage points. Kemptner et al. (2011) find in their IV-specification a reduction of 3.1 percentage points in the overweight probability, we find 7.5 percentage points for the respective specification. However, we also find significant effects on obesity in the Jürges et al. (2009) specification.

5.2 Model specification

Table 6 provides the results for several sensitivity checks regarding the model specification. First, we include children of mothers without siblings (specification 7) to show that our findings are robust even when the functional form is presumably misspecified.[10] Second, we estimate a log specification where maternal education is instrumented by the natural logarithm of the number of her siblings (specification 8). Third, allowing for full flexibility with respect to the functional relationship between the number of siblings and maternal education, we instrument maternal education with a set of dichotomous variables indicating the number of siblings (specification 9). Categories of the number of siblings are 1, 2, 3, 4, 5 and 6 and more siblings. Fourth, we check the sensitivity of our findings regarding the distributional assumptions of the IV-Probit model (specification 10). The baseline specification assumes that the error terms of equation (2) and (3) are bivariate normally and identically distributed while allowing for serial correlation between children with the same mother. Both distributional assumptions can be relaxed by estimating a simple two-stage least squares model (2SLS) with cluster-robust standard errors. Unlike the IV-Probit model, the 2SLS model also produces consistent parameter estimates in the presence of heteroscedasticity and non-normally distributed errors.

In all these additional specifications, the estimated effects change only marginally. In general, when including mothers without siblings, estimating the log specification or using the 2SLS estimator, the estimated coefficients exceed the coefficients in the baseline specification. Hence, our baseline specification is more conservative. Aside from the effect on doing sports regularly for daughters in specification 9, all effects remain significant. This insignificance is due to the loss in precision in the first stage when using a set of dummy variables as instruments. Thus, we conclude that neither the specification of the functional relationship between number of siblings and education nor the distributional assumptions

[10] Mothers without siblings seem to be special. They have, on average, less education than mothers with one or two siblings (see the discussion in section 3).

regarding the error terms of the IV-Probit model compromise the consistency of our estimates.

6. Channels

This subsection discusses two possible channels of causality that could drive the causal relationship between maternal education and the child's outcomes. Table 7 contains the results of two alternative specifications and the baseline specification. Again, we only consider outcome variables with significant effects in the baseline specification.

Assortative mating may explain to some degree the effects of maternal education on child outcomes. Mothers are very likely to have a partner with a similar level of education. In our data we find a correlation coefficient of 0.66 between maternal years of education and her partner's years of education. Thus, the estimated causal effects may work through the partner's education. We focus on the mother's *partner* because the SOEP does not report on the biological father but only on the mother's current partner. In specification 11, we estimate effects of maternal education on the child's outcomes, including the partner's years of education as an additional control variable. The results must be interpreted carefully as the partner's education is likely to be an endogenous variable. This may also bias the estimates for the effects of maternal education. Furthermore, the number of cases drop because we only consider mothers with a partner.

The magnitude of the effects on the probabilities of the daughter being a smoker at age 18/19 and of the daughter doing sports regularly at age 18/19 change only marginally, but the former effect becomes insignificant. This may be due to the substantial loss in precision. The effect on the likelihood of being on an academic track school at age 17 becomes small (the sign even becomes negative) and insignificant for sons. The effect for daughters remains large and significant. The effect on the probability of repeating a grade remains large and significant for both sons and daughters. Thus, our findings suggest that assortative mating of the mother is relevant. But, it seems that the mother's partner is more relevant for sons than for daughters. Assortative mating appears to be most important with respect to sons' schooling track. Mothers appear to be more influential with respect to their daughters. Our estimates become relatively imprecise when controlling for assortative mating.[11]

The effects of maternal education on the child's outcomes may also work through a higher household income. We investigate this issue including the logarithm of a five years average of household post-government income as an additional control variable (specification 12).[12] To some extent this accounts for assortative mating because the mothers' partners are the principle earners in the majority of the households. Again, the results have to be interpreted cautiously because also the household income is likely to be an endogenous regressor. With this specification the estimated effects change only slightly. The effect on the daughter doing sports regularly becomes insignificant. Overall, we infer that household income does not explain the estimated effects of maternal education.

7. Summary and discussion

This paper is the first to investigate the causal effect of maternal education on health related behaviour and schooling outcomes of children in Germany. Using a rich survey panel data set, we estimate the causal effect on a wide range of outcomes for adolescent children. We find strong and significant effects on schooling outcomes for both sexes and also on health-related behaviour for daughters. A unit increase in maternal years of education is estimated to reduce the child's probability to repeat a grade by 7.4 percentage points, and to increase the child's probability of pursuing an academic schooling track by 9.5 percentage points. The findings on schooling outcomes do not differ by the child's sex. The magnitude of our estimated effects

[11] We also estimated a model instrumenting both the education of the mother and of the partner, each by the number of siblings (not displayed). This specification is even more imprecise due to fewer observations: We needed information on all four grandparents of the adolescent and had to exclude mothers and partners without siblings. The results correspond qualitatively to specification (11).
[12] In the case that household income is missing for some years, we only use the available years for mean computation. All annual household incomes are transformed to 2006 Euros.

on grade repetition is in the range of findings for the U.S. (e.g. Carneiro et al. 2007, Oreopoulos et al. 2003).

Furthermore, we estimate that an increase in maternal education by one year reduces the adolescent daughter's probability to smoke at age 18/19 by 5.8 percentage points and increases the adolescent daughter's likelihood of doing sports at least once a week by 6.7 percentage points. However, we do not obtain significant effects of maternal education on sons' health behaviour.

In line with previous research (e.g. Carneiro et al. 2007, Currie/Moretti 2003, Oreopoulos et al. 2003) our IV-Probit estimates exceed the corresponding estimates from the Probit model. This might be attributed to three different reasons: measurement error in maternal education, unobserved variables leading to downward biased estimates in the Probit model, or the identification of local effects in the presence of effect heterogeneity when applying an IV-approach.

For our identification strategy, we do not rely on policy changes like previous studies. Instead, we present an IV approach that also works for cohorts unaffected by policy changes and for the limited sample sizes of common household panels. For this purpose, we argue that the mother's number of siblings is - conditional on the social status of her parents - a valid and relevant instrument for maternal education. Concerning the relevance of the instrument, we find all respective first stage F-statistics to exceed the critical value of 10. The estimation strategy is not suffering from a weak instruments problem. To underline the validity of the instrument, we show that our results are robust over a variety of different specifications. Including more detailed measures of the grandparents' social status, the number of the mother's children or indicators of mother's own health and health-related behaviour does not alter the results. The exclusion of mothers who grew up in large cities from our sample does not change the findings substantially. Furthermore, we present evidence that neither the specification of the functional relationship between the number of siblings and maternal education nor the distributional assumptions regarding the error terms of the IV-Probit model compromise the consistency of our estimates.

Investigating possible channels of causality, our findings suggest that assortative mating of the mother is relevant. But, it seems that the mother's partner is more relevant for sons than for daughters. Assortative mating appears to be most important with respect to sons' schooling track. Together with the fact that we only find effects on the health behaviour of daughters, this suggests that mothers are more influential regarding their daughters.

We find substantial intergenerational spillover effects of maternal education. The number of mother's siblings constrains her educational investments and affects also her children's human capital. Public policy should take into account these intergenerational links when thinking about optimal educational investments. There are persistent gains to be realized by increasing female education. The reduction of financial constraints on educational investments could be a means to this end. This is also a matter of social mobility within the German society.

8. References

Andersen, Hanfried H., Axel Mühlbacher, Matthias Nübling, Jürgen Schupp, and Gert G. Wagner. 2007. Computation of standard values for physical and mental health scale scores using the SOEP version of SF-12v2. *Schmollers Jahrbuch* 127: 171-182.

Björklund, Anders, Mikael Lindahl, and Erik Plug. 2006. The origins of intergenerational associations: Lessons from Swedish adoption data. *Quarterly Journal of Economics* 121: 999-1028.

Black, Sandra E., and Paul J. Devereux. 2010. Recent developments in intergenerational mobility. *NBER Working Paper* 4866.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2005. The more the merrier? The effect of family size and birth order on children's education. *Quarterly Journal of Economics* 120: 669-700.

Card, David. 1999. The causal effect of education on earnings. In *Handbook of Labor Economics 3A*, vol. 3, ed. Orley Ashenfelter. New York: Elsevier.

Carneiro, Pedro, Costas Meghir, and Matthias Parey. 2007. Maternal education, home environments and the development of children and adolescents. *IZA Discussion Paper* 3072.

Coneus, Katja, and C. Katharina Spieß. 2008. The Intergenerational transmission of health in early childhood. *SOEPpapers on Multidisciplinary Panel Data Research* 126.

Currie, Janet. 2009. Healthy, wealthy, and wise: Socioeconomic status, poor health in childhood, and human capital development. *Journal of Economic Literature* 47: 87-122.

Currie, Janet, and Enrico Moretti. 2003. Mother's education and the intergenerational transmission of human capital: Evidence from college openings. *Quarterly Journal of Economics* 118: 1495-1532.

Cutler, David M., and Adriana Lleras-Muney. 2008. Education and health: evaluating theories and evidence. P. 29–60 in *Making Americans Healthier: Social and Economic Policy as Health Policy*, edited by Robert F. Schoeni, James S. House, George A. Kaplan, and Harold Pollack. New York: Russell Sage Foundation.

Ganzeboom, Harry B.G., Paul M. de Graaf, and Donald J. Treiman. 1992. A standard international socio-economic index of occupational status. *Social Science Research* 21: 1-56.

Gebel, Michael, and Friedhelm Pfeiffer. 2010. Educational expansion and its heterogeneous returns for wage workers. *Schmollers Jahrbuch* 130: 19-42.

Grossman, Michael. 2006. Education and nonmarket outcomes. P. 577-633 in *Handbook of the Economics of Education*, vol. 1, edited by Eric A. Hanushek and F.Welch. Amsterdam: Elsevier.

Heineck, Guido, and Regina T. Riphahn. 2007. Intergenerational transmission of educational attainment in Germany: The last five decades. *IZA Discussion Paper* 2985.

Holmlund, Helena, Mikael Lindahl, and Erik Plug. 2008. The causal effect of parent's schooling on children's schooling: a comparison of estimation methods. *IZA Discussion Paper* 3630.

Jackson, Gregg B. 1975. The research evidence on the effects of grade retention. *Review of Educational Research* 45: 613-635.

Jürges, Hendrik, Steffen Reinhold, and Martin Salm. 2009. Does schooling affect health behavior? Evidence from the educational expansion in Western Germany. *IZA Discussion Paper* 4330.

Kemptner, Daniel, Hendrik Jürges, and Steffen Reinhold. 2011. Changes in Compulsory Schooling and the Causal Effect of Education on Health: Evidence from Germany. *Journal of Health Economics* 30: 340-354.

Lamerz, A., J. Kuepper-Nybelen, C. Wehle, N. Bruning, G. Trost-Brinkhues, H. Brenner, J. Hebebrand, and B. Herpertz-Dahlmann. 2005. Social class, parental education, and obesity prevalence in a study of six-year-old children in Germany. *International Journal of Obesity* 29: 373-380.

Lindeboom, Maarten, Ana Llena-Nozal, and Bas van der Klaauw. 2009. Parental education and child health: evidence from a schooling reform. *Journal of Health Economics* 28: 109-131.

Little, Roderick J., and Donald B. Rubin. 2002. *Statistical analysis with missing data*. 2nd ed. Hoboken, NJ: Wiley.

Loureiro, Maria L., Anna Sanz-de-Galdeano, and Daniela Vuri. 2010. Smoking Habits: Like Father, Like Son, Like Mother, Like Daughter?. *Oxford Bulletin of Economics and Statistics* 72: 717-743.

Maurin, Eric, and Sandra McNally. 2008. Vive la révolution! Long term educational returns of 1968 to the angry students. *Journal of Labor Economics* 26: 1-33.

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

Oreopoulos, Philip, Marianne E. Page, and Ann H. Stevens. 2003. Does human capital transfer from parent to child? The intergenerational effects of compulsory schooling. *NBER Working Paper* 10164.

Oreopoulos, Philip, Marianne E. Page, and Ann H. Stevens. 2006. The intergenerational effects of compulsory schooling. *Journal of Labor Economics* 24: 729-760.

Pischke, Jörn-Steffen, and Till von Wachter. 2008. Zero Returns to Compulsory Schooling in Germany: Evidence and Interpretation. *Review of Economics and Statistics* 90: 592-598.

Reinhold, Steffen, and Hendrik Jürges. 2010. Secondary school fees and the causal effect of schooling on health behaviour. *Health Economics* 19: 994-1001.

Rubin, Donald B. 1976. Inference and missing data. *Biometrika* 63: 581-592.

Sander, William. 1995. Schooling and quitting smoking. *Review of Economics and Statistics* 77: 191-199.

Starrels, Marjorie E. 1994. Gender differences in parent-child relations. *Journal of Family Issues* 15: 148-165.

Wagner, Gert G., Joachim R. Frick, and Jürgen Schupp. 2007. The German socio-economic panel study (SOEP) - Scope, evolution and enhancements. *Schmollers Jahrbuch* 127: 139-169.

A Appendix

Some variables on the mothers' and grandparents' level are affected by missing values; e.g. 37% of the children in the newborn sample have missing information on either the area where the mother grew up, grandparents' ISEI score, or grandfather's and grandmother's educational level. Not considering these cases will produce inefficient estimates, even if they are missing completely at random (MCAR; see Rubin 1976). The estimates will be biased if the information is not MCAR but only missing at random (MAR). Under MAR the missingness depends on other observed variables, e.g. if mothers with fewer years of education know less about their parents.

Due to these effectiveness and unbiasedness considerations, we impute four variables relevant for our analysis:[13] grandfather's and grandmother's educational level, grandparents' ISEI score and the area where the mother grew up. For all variables we first copy consistent information provided by the mother's siblings in case the information is missing. We impute missing values in the size of mother's area randomly conditional on the size of the mother's district of residence when she was interviewed for the first time.

The other three variables are jointly imputed in four steps as follows. First, the educational levels of the grandparents are preliminarily imputed: If the level of education is missing for only one grandparent the information of the other grandparent is used. If the level of vocational training is available, the mode of level of education for each vocational training category is imputed. Second, we run a regression of the highest ISEI score of the grandparents (in most cases the grandfather's) score on sets of dummies for the grandfather's levels of vocational training and education, as well as dummies for the grandmother's levels of education and vocational training, dummies for the job position[14] of the grandfather, controls

[13] The imputation procedure uses information on all SOEP respondents with at least one personal interview with biography information. The terms grandfather and grandmother are, hence, incorrect in a strict sense but merely applied to describe the generation.

[14] This variable is not about the exact occupation of the grandfather but rather a general description of his class of job, e.g. blue collar, agricultural worker, self-employed.

for the birth decade of the grandfather and for each explanatory variable a dummy for missing values. These variables explain about 2/3 of the variance in grandparents' ISEI score. We exclude observations with missing information on all explanatory variables and do not impute any values for them.

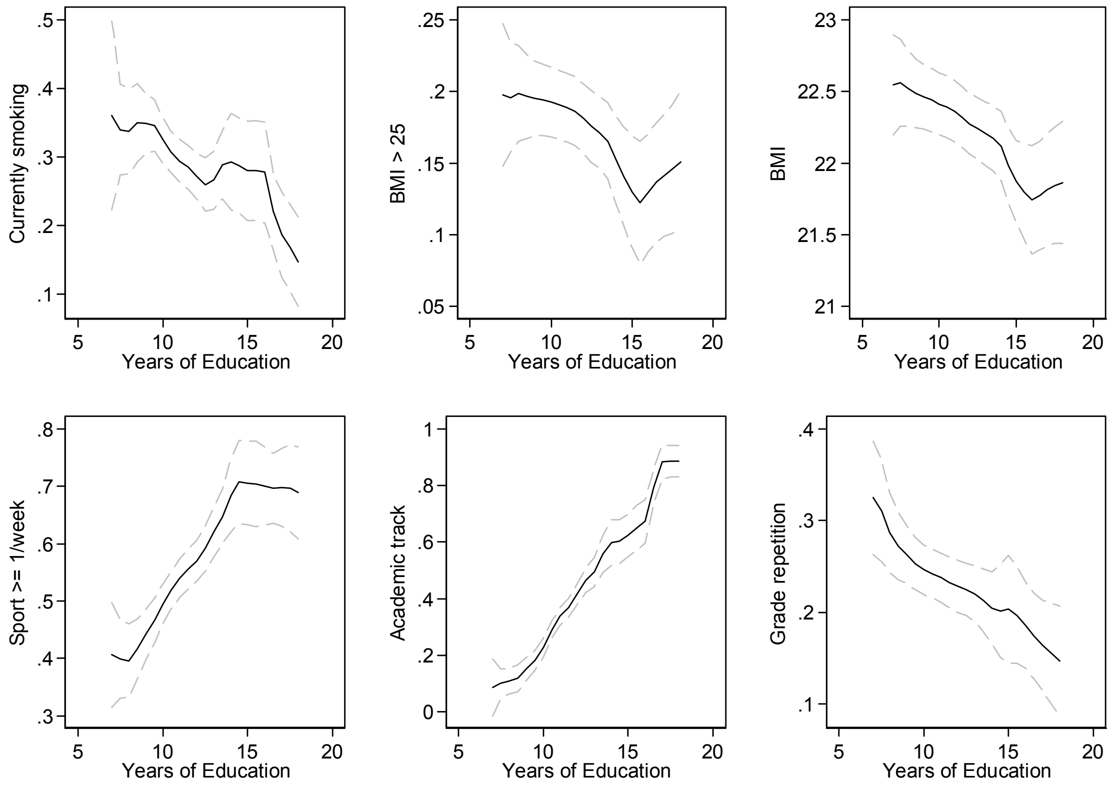
Third, according to the regression results we predict values for those with missing information on the grandparents' ISEI score. We then add a random term drawn from the distribution of the regression residuals to maintain the variance of the dependent variable and to mimic the uncertainty of the imputation. Little & Rubin (2002: 60) refer to this procedure as stochastic regression imputation. Fourth, by means of multinomial logit models we regress the grandparents' educational level on the imputed grandparents' ISEI score, dummies for own vocational training levels and partner's education level. We use the predicted level of education for all those with missing information, including those with preliminarily imputed educational levels.

Table 1: Descriptive Statistics

Generation	Variable	All mothers		Mothers with siblings	
		Mean	**SD**	**Mean**	**SD**
Mother	Year of birth	1959.0	(5.33)	1959.2	(5.28)
	Years of education	11.76	(2.62)	11.75	(2.64)
	Number of siblings	2.27	(1.98)	2.65	(1.89)
Child	Year of birth	1986.5	(2.28)	1986.6	(2.27)
	Smoker (yes/no)	0.284	(0.451)	0.292	0.455
	Overweight (yes/no)	0.174	(0.379)	0.178	0.383
	Sport (yes/no)	0.550	(0.498)	0.553	0.497
	Pursuing *Abitur* (yes/no)	0.393	(0.489)	0.385	0.487
	Grade repetition (yes/no)	0.239	(0.426)	0.243	0.429
Number of Observations		1852		1590	

Note: Unweighted means and standard deviations for key variables of children and their mothers. The last two columns exclude mothers without siblings.

Figure 1: The bivariate relationship



Note: The bivariate relationship between maternal education and various child outcomes. The lines picture local mean smooths computed with Stata's default bandwidth and an Epanechnikov kernel, as well as the 95% confidence bands.

Table 2: Effect of number of mother's siblings on maternal education

Number of siblings fixed effects	No demographic controls	Demographic controls
1 sibling	0.654***	0.588***
	(0.195)	(0.169)
2 siblings	0.214	0.176
	(0.198)	(0.171)
3 siblings	-0.481**	-0.196
	(0.226)	(0.196)
4 siblings	-0.449*	0.015
	(0.256)	(0.223)
5 siblings	-0.803**	-0.489*
	(0.328)	(0.283)
6 siblings	-1.674***	-0.806**
	(0.394)	(0.343)
7 siblings	-1.319***	-0.798*
	(0.507)	(0.438)
8 siblings	-1.284**	-1.140**
	(0.606)	(0.523)
9 siblings	-2.459***	-1.331***
	(0.592)	(0.516)
10+ siblings	-2.130***	-1.188*
	(0.725)	(0.629)
Constant	11.784***	10.734***
	(0.158)	(0.336)
Number of Observations	1839	

Note: Marginal effects and robust standard errors (in parenthesis) for the sample of mothers. The regression with demographic controls contains controls for the child's sex, the child's year of birth, the mother's age at birth of the child, fixed effects for the area the mother grew up and for the educational levels of the mother's parents. * p<10 %, ** p<5 %, *** p<1 %.

Table 3: First stage results for IV-Probit models

Dep. Var.	**Sex**	**Average Marginal Effects First Stage IV-Probit**		**Obs.**
		Spec. 1 [minimal]	**Spec. 2 [baseline]**	
Smoker	pooled	-0.334***	-0.214***	1365
		(0.033)	(0.031)	
	sons	-0.306***	-0.174***	682
		(0.047)	(0.045)	
	daughters	-0.377***	-0.267***	683
		(0.047)	(0.044)	
Overweight	pooled	-0.331***	-0.208***	1338
		(0.034)	(0.032)	
	sons	-0.303***	-0.162***	664
		(0.048)	(0.046)	
	daughters	-0.378***	-0.263***	674
		(0.049)	(0.045)	
Sport	pooled	-0.346***	-0.232***	1345
		(0.034)	(0.031)	
	sons	-0.339***	-0.208***	686
		(0.047)	(0.045)	
	daughters	-0.366***	-0.268***	659
		(0.049)	(0.046)	
Academic track	pooled	-0.341***	-0.225***	1440
		(0.034)	(0.032)	
	sons	-0.312***	-0.195***	727
		(0.048)	(0.046)	
	daughters	-0.390***	-0.278***	713
		(0.051)	(0.048)	
Grade repetition	pooled	-0.338***	-0.227***	1506
		(0.033)	(0.031)	
	sons	-0.320***	-0.199***	761
		(0.046)	(0.044)	
	daughters	-0.372***	-0.270***	745
		(0.048)	(0.045)	

Note: Average marginal effects of the number of mother's siblings on her years of education and robust standard errors (in parenthesis) for first stage of IV-Probit models (pooled and gender-specific). All regressions include controls for the child's sex (if pooled sample) as well as for the child's and the mother's year of birth. The baseline specification (2) includes additional fixed effects for the area the mother grew up and for the educational levels of the mother's parents. * p<10 %, ** p<5 %, *** p<1 %.

Table 4: Results for Probit and IV-Probit models

Dep. Var.	Sex	Average Marginal Effects			
		Second Stage IV-Probit		Probit	Obs.
		Spec. 1 [minimal]	Spec. 2 [baseline]	Spec. 2 [baseline]	
Smoker	pooled	-0.018	-0.019	-0.017***	1365
		(0.019)	(0.031)	(0.005)	
	sons	0.011	0.045	-0.013*	682
		(0.030)	(0.047)	(0.008)	
	daughters	-0.037*	-0.058**	-0.021***	683
		(0.021)	(0.029)	(0.008)	
Overweight	pooled	-0.010	-0.011	-0.010**	1338
		(0.016)	(0.027)	(0.005)	
	sons	-0.038	-0.059	-0.013	664
		(0.025)	(0.047)	(0.008)	
	daughters	0.018	0.025	-0.008	674
		(0.020)	(0.030)	(0.007)	
Sport	pooled	0.043**	0.053**	0.034***	1345
		(0.018)	(0.027)	(0.006)	
	sons	0.017	0.029	0.034***	686
		(0.027)	(0.046)	(0.008)	
	daughters	0.067***	0.067**	0.034***	659
		(0.019)	(0.032)	(0.008)	
Academic track	pooled	0.094***	0.095***	0.068***	1440
		(0.009)	(0.016)	(0.004)	
	sons	0.091***	0.090***	0.070***	727
		(0.015)	(0.029)	(0.006)	
	daughters	0.097***	0.103***	0.067***	713
		(0.010)	(0.016)	(0.007)	
Grade repetition	pooled	-0.055***	-0.074***	-0.018***	1506
		(0.013)	(0.019)	(0.005)	
	sons	-0.063***	-0.083***	-0.016**	761
		(0.019)	(0.028)	(0.007)	
	daughters	-0.048***	-0.070***	-0.019***	745
		(0.018)	(0.024)	(0.007)	

Note: Average marginal effects of maternal education and robust standard errors (in parenthesis) for second stage of IV- Probit models (pooled and gender-specific). All regressions include controls for the child's sex (if pooled sample) as well as for the child's and the mother's year of birth. The baseline specification (2) includes additional fixed effects for the area the mother grew up and for the educational levels of the mother's parents. * p<10 %, ** p<5 %, *** p<1 %.

Table 5: Sensitivity checks – validity assumption

Dep. Var.	Sex	Average Marginal Effects Second Stage IV-Probit			
		Spec. 3 [social status]	Spec. 4 [mother's health]	Spec. 5 [large cities]	Spec. 6 [fertility]
Smoker	pooled	-0.029	-0.023	-0.019	-0.018
		(0.038)	(0.034)	(0.035)	(0.033)
		1365	1328	1096	1352
	sons	0.054	0.050	0.044	0.043
		(0.056)	(0.053)	(0.054)	(0.053)
		682	662	554	676
	daughters	-0.071**	-0.061**	-0.052	-0.057*
		(0.031)	(0.030)	(0.032)	(0.031)
		683	666	542	676
Sport	pooled	0.046	0.044	0.062**	0.059**
		(0.035)	(0.034)	(0.028)	(0.027)
		1345	1273	1072	1332
	sons	0.029	-0.004	0.061	0.021
		(0.059)	(0.061)	(0.045)	(0.052)
		686	649	557	679
	daughters	0.060	0.070**	0.058	0.083***
		(0.039)	(0.034)	(0.036)	(0.025)
		659	624	515	653
Academic track	pooled	0.096***	0.093***	0.108***	0.090***
		(0.021)	(0.020)	(0.016)	(0.019)
		1440	1306	1150	1427
	sons	0.084**	0.087**	0.120***	0.082**
		(0.041)	(0.036)	(0.025)	(0.036)
		727	662	591	721
	daughters	0.106***	0.104***	0.104***	0.100***
		(0.019)	(0.018)	(0.018)	(0.019)
		713	644	559	706
Grade repetition	pooled	-0.085***	-0.079***	-0.068***	-0.076***
		(0.021)	(0.022)	(0.024)	(0.020)
		1506	1361	1203	1493
	sons	-0.096***	-0.090***	-0.099***	-0.088***
		(0.029)	(0.031)	(0.029)	(0.029)
		761	690	617	755
	daughters	-0.073***	-0.079***	-0.040	-0.068***
		(0.028)	(0.026)	(0.032)	(0.025)
		745	671	586	738

Note: Average marginal effects of maternal education, robust standard errors (in parenthesis) and number of observations for second stage of IV-Probit models (pooled and gender-specific). All regressions include controls for the child's sex (if pooled sample), the child's and the mother's year of birth as well as fixed effects for the area the mother grew up and for the educational levels of the mother's parents. Specification (3) controls additionally for the log of the grandfather's ISEI score and a migration dummy. Specification (4) includes in addition to (2) controls for mother's BMI, measures of her physical and mental health, as well as a variable indicating the mother being a smoker. Specification (5) excludes mothers who grew up in large cities. (6) includes additional dummies for 1, 2, 3 and 4 or more children of the mother. * p<10 %, ** p<5 %, *** p<1 %.

Table 6: Sensitivity checks – model specification

Dep. Var.	Sex	Average Marginal Effects Second Stage IV-Probit			
		Spec. 7 [+only child]	Spec. 8 [log]	Spec. 9 [dummies]	Spec. 10 [2SLS]
Smoker	pooled	-0.060*	-0.034	-0.017	-0.016
		(0.032)	(0.029)	(0.033)	(0.033)
		1578	1365	1365	1365
	sons	-0.022	0.028	0.045	0.053
		(0.067)	(0.047)	(0.052)	(0.059)
		782	682	682	682
	daughters	-0.080**	-0.075***	-0.062**	-0.061*
		(0.032)	(0.027)	(0.030)	(0.037)
		796	683	683	683
Sport	pooled	0.040	0.058**	0.060**	0.056*
		(0.037)	(0.026)	(0.026)	(0.032)
		1564	1345	1345	1345
	sons	-0.011	0.018	0.050	0.029
		(0.058)	(0.046)	(0.052)	(0.048)
		784	686	686	686
	daughters	0.075**	0.085***	0.065	0.074*
		(0.036)	(0.024)	(0.042)	(0.043)
		780	659	659	659
Academic track	pooled	0.111***	0.095***	0.099***	0.104***
		(0.014)	(0.015)	(0.014)	(0.028)
		1674	1440	1440	1440
	sons	0.121***	0.097***	0.100***	0.100**
		(0.016)	(0.022)	(0.023)	(0.047)
		827	727	727	727
	daughters	0.106***	0.100***	0.101***	0.111***
		(0.022)	(0.018)	(0.016)	(0.032)
		847	713	713	713
Grade repetition	pooled	-0.087***	-0.088***	-0.083**	-0.083***
		(0.021)	(0.016)	(0.033)	(0.029)
		1752	1506	1506	1506
	sons	-0.106***	-0.095***	-0.106***	-0.099**
		(0.022)	(0.022)	(0.028)	(0.050)
		873	761	761	761
	daughters	-0.069**	-0.087***	-0.061	-0.073**
		(0.034)	(0.022)	(0.046)	(0.034)
		879	745	745	745

Note: Average marginal effects of maternal education, robust standard errors (in parenthesis) and number of observations for second stage of IV-Probit models (pooled and gender-specific). All regressions include controls for the child's sex (if pooled sample), the child's and the mother's year of birth as well as fixed effects for the area the mother grew up and for the educational levels of the mother's parents. Specification (7) includes also mothers without siblings. Specification (8) instruments maternal years of education with the natural *logarithm* of the number of siblings, while specification (9) rests upon dummies for 1, 2, 3, 4, 5 and 6+ siblings as instruments. Specification (10) is estimated with two-stage least squares. * p<10 %, ** p<5 %, *** p<1 %.

Table 7: Channels – partner's education and household income

Dep. Var.	Sex	Average Marginal Effects Second Stage IV-Probit		
		Spec. 11 [mating]	**Spec. 12 [income]**	**Spec. 2 [baseline]**
Smoker	pooled	-0.007	-0.013	-0.019
		(0.050)	(0.040)	(0.031)
		1096	1365	1365
	sons	0.082	0.062	0.045
		(0.081)	(0.050)	(0.047)
		541	682	682
	daughters	-0.049	-0.066*	-0.058**
		(0.042)	(0.038)	(0.029)
		555	683	683
Sport	pooled	0.050	0.043	0.053**
		(0.055)	(0.038)	(0.027)
		1043	1345	1345
	sons	-0.025	0.009	0.029
		(0.125)	(0.062)	(0.046)
		528	686	686
	daughters	0.074*	0.064	0.067**
		(0.045)	(0.044)	(0.032)
		515	659	659
Academic track	pooled	0.056	0.090***	0.095***
		(0.045)	(0.024)	(0.016)
		1072	1440	1440
	sons	-0.022	0.077*	0.090***
		(0.099)	(0.046)	(0.029)
		539	727	727
	daughters	0.089***	0.104***	0.103***
		(0.032)	(0.022)	(0.016)
		533	713	713
Grade repetition	pooled	-0.093***	-0.083***	-0.074***
		(0.029)	(0.023)	(0.019)
		1117	1506	1506
	sons	-0.121***	-0.089***	-0.083***
		(0.041)	(0.034)	(0.028)
		562	761	761
	daughters	-0.072**	-0.080***	-0.070***
		(0.035)	(0.029)	(0.024)
		547	745	745

Note: Average marginal effects of maternal education, robust standard errors (in parenthesis) and number of observations for second stage of IV-Probit models of the sample of adolescents (pooled and gender-specific). All regressions include controls for the child's sex (if pooled sample), the child's and the mother's year of birth as well as fixed effects for the area the mother grew up and for the educational levels of the mother's parents. Specification (7) controls for the years of education of the mother's partner, specification (8) includes the logarithm of a five years average of household income. * p<10 %, ** p<5 %, *** p<1 %.