Hörisch, Hannah

Working Paper

Does parental employment affect children's educational attainment? Evidence from Germany

Munich Discussion Paper, No. 2008-5

Provided in Cooperation with:
University of Munich, Department of Economics

Suggested Citation: Hörisch, Hannah (2008) : Does parental employment affect children's educational attainment? Evidence from Germany, Munich Discussion Paper, No. 2008-5, Ludwig-Maximilians-Universität München, Volkswirtschaftliche Fakultät, München, https://doi.org/10.5282/ubm/epub.2140

This Version is available at:
https://hdl.handle.net/10419/104264

Hannah Hörisch:

Does parental employment affect children's educational attainment?

Munich Discussion Paper No. 2008-5

Department of Economics
University of Munich

Volkswirtschaftliche Fakultät
Ludwig-Maximilians-Universität München

Online at http://epub.ub.uni-muenchen.de/2140/

Does parental employment affect children's educational attainment? Evidence from Germany*

Hannah Hörisch
University of Munich†

February 26, 2008

Abstract

This paper analyzes whether there exists a causal relationship between parental employment and children's educational attainment. We address potential endogeneity problems due to (i) selection of parents in the labor market by estimating a model on sibling differences and (ii) reverse causality by focusing on parents' employment when children are aged 0-3. We use data from the German Socioeconomic Panel. Overall, we find little support that parental employment affects children's educational attainment. We can rule out that having a mother who works one hour more per week lowers the probability of high secondary track attendance by more than 0.1%.

1 Introduction

Over the last decades, female labor market participation rates and especially those of mothers with young children have increased tremendously in many countries. In the US, 47 % of mothers with children below age 6 worked in 1975. By 2006, this share had increased to 71 % (Chao and Rones, 2007, Table 7). In Germany, 35 % of mothers with children below age 6 worked in 1974, but 52 % in 2004. In contrast,

*I thank Julia Bersch, Nadine Riedel, Klaus Schmidt, Guido Schwerdt, Joachim Winter, Ludger Wößmann for helpful comments and discussions. Financial support from the Deutsche Forschungsgemeinschaft through SFB/TR 15 is gratefully acknowledged.

†*Affiliation:* University of Munich, Seminar for Economic Theory, Ludwigstraße 28 (Rgb.), 80539 Munich, Germany, email: hannah.hoerisch[at]lrz.uni-muenchen.de

labor market participation rates of German fathers have remained very stable at about 88 %.[1]

Precise knowledge about how parental employment affects children's long-term outcomes such as educational attainments or labor market success is crucial for the evaluation of many policy programs. For example, US welfare reforms in the 1990s pushed welfare recipients and in particular welfare dependent single mothers to find employment (compare Blank, 2002). Reforms were motivated by the belief that parental work is the best way out of poverty for parents and children. If, however, having working parents hurts the educational and labor market prospects of children such reforms may be counterproductive in the long run. To give another example, the current German government's decision to substantially expand and subsidize day care facilities for children below age 3 has lead to emotional and controversial debates in the German public. Opponents of day care expansion consider full-time parental child care to be decisive for children's cognitive and emotional development. Proponents argue that parent-child interactions can be substituted by high quality non-parental child care and that increases in family income may also benefit children.

This paper is the first to use a large German household panel data set, the German Socioeconomic Panel (GSOEP), to analyze whether parental employment hurts or benefits children's educational attainments. Our measure of educational attainment and dependent variable is attendance or completion of high secondary school track (so called Gymnasium) which is the only track that provides direct access to university. We separately analyze two effects of parental employment: first, the effect on income that may influence child-related investments, i.e. we control for total household income. Second, we use three different measures of parental time inputs in raising their children to capture the "time effect" of parental employment: weekly hours worked, the number of years in which parents work full-time, part-time or not at all and weekly hours that parents spent on child care when children are aged 0-3.

We explicitly approach potential endogeneity problems. First, to take selection of parents in the labor market into account we estimate a model on sibling differences that controls for all unobserved time-invariant parent and household characteristics. Second, we address the potential reverse causality problem, i.e. the fact that parents' decisions to work may be influenced by their child's ability which in turn affects educational attainment. We focus on parental employment when children are aged

[1] Figures stem from an inquiry at the Federal Statistical Office of Germany. The 35 % in 1974 refer to West Germany only.

0-3 such that a child's ability is not yet fully revealed, exclude disabled children from the analysis and use parents' years of education as a proxy for their child's ability.

We do not find any evidence that parental employment hurts children's educational attainment. Controlling for household income we can statistically rule out that having a mother who works one hour more per week lowers the probability of high secondary track attendance by more than 0.1 percentage points. Actually, all coefficients of maternal employment are positive, but not significant at conventional levels (though at a 9 to 11 % significance level). Coefficients of fathers' employment and parental time spent on child care are precisely estimated, but too small to be significant. Testing for equality of mother's and father's time input coefficients, we cannot reject that parents' time inputs are substitutes.

Table 1 reviews results from previous economic studies that investigate the relationship between parental employment and children's educational attainment. In sum, evidence is very inconclusive: some studies predict a negative effect of parental employment on children's educational attainment, some a positive one and the remainder insignificant effects or effects that differ by subsamples such as sex or race of the child. Table 1 also reveals some characterizing features of existing studies. First and most importantly, except for Ermisch and Francesconi (2002) the studies listed in Table 1 ignore problems that arise due to omitted variables such as a child's ability or selection of parents into the labor market. In contrast, this paper addresses the conditions under which we obtain consistent estimates explicitly and estimates a model on sibling differences to control for unobserved parent and household characteristics. Second, only two studies (Ermisch and Francesconi, 2002 and O'Brien and Jones, 1999) report estimates on the effect of father's employment. Our paper estimates the effects of parental employment separately for mothers and fathers and as Ermisch and Francesconi (2002) also the joint effect of e.g. hours worked. Third, all studies use US or British data. Since the institutional environment (child care facilities, maternity leave policies, etc.) and the attitudes towards working mothers differ substantially across countries, evidence from Anglo-Saxon countries might not be transferable to other Western countries. Our study adds evidence from Germany to the existing literature. Last, all studies in Table 1 use indirect measures of parental time inputs such as the type of parental employment (full-time, part-time or none) or years worked.[2] An advantage of the GSOEP data

[2]Haveman, Wolfe and Spaulding (1991) use *estimated* parental time spent on child care as explanatory variable. They do not have information on time spent on child care in their original data but construct it from a second data set. Using the second data set they regress child care

Table 1: Related literature

study	data source, country	outcome	estimation method	effect of parental employment*
Ermisch and Francesconi (2002)	British Household Panel Survey, UK	highest educational qualification (A level or more)	logit, linear probability models, sibling differences model	mother works part-time: level estimates: (-) ns sibling difference est.: (-) 10 mother works full-time: level estimates: (-) ns sibling difference est.: (-) 5 father works: level estimates: (+) 5 sibling difference est.: (-) ns
Graham, Beller and Hernandez (1994)	Current Population Survey, US	years of schooling at ages 16-20	2SLS, first stage: IV for child support	mother worked outside home: (+) 1
Haveman, Wolfe and Spaulding (1991)	Panel Study of Income Dynamics, US	high school graduation	probit model	years mother worked: (+) 1
Hill and Duncan (1987)	Panel Study of Income Dynamics, US	years of schooling at ages 27-29	OLS, gender specific	mother's work hours: for men: (-) 5 for women: (-) ns
Kiernan (1996)	National Child Development Study, UK	no degree	descriptive statistics, logit model	mother's non-employment: for men: no effect for women: (+) 1
Krein and Beller (1988)	National Longitudinal Surveys, US	years of schooling at age 26	OLS, gender and race specific	mother ever worked outside home at least 6 months at ages 0-18: white men: (-) 1 white women: (-) ns black men: (+) ns black women: (-) ns
O'Brien and Jones (1999)	survey and time-use diaries in 6 schools in East London, UK	highest / lowest national test scores	logit model	low educational outcome: father works full time and mother full time: (-) ns mother part-time: (-) 5 high educational outcome: father works full-time and mother full time: (+) ns mother part-time: (+) 10

* (-) indicates a negative sign, (+) a positive sign

ns: not significant; 1, 5, 10: significant at a 1, 5, 10 percent significance level

set is that it contains very detailed information on the time parents spent on child care. Besides the commonly used indirect measures we use the hours parents spent on child care on a typical weekday when children are aged 0-3 as a direct measure of parental time inputs in raising their children.

Haveman and Wolfe (1995) review studies on the effects of parental employment on a broad range of children's outcomes such as high school graduation, years of schooling, out-of-wedlock fertility or adult earnings. All reviewed studies use US data and do not address endogeneity problems. Ermisch and Francesconi (2005) survey the more recent literature on parental employment and children's well-being covering studies that use data from countries different from the US, mainly from UK.

Using German administrative data Dustmann and Schönberg (2007) analyze the effect of three extensions in maternity leave coverage on children's later attendance of high secondary track and wages. They compare cohorts of children born shortly before and after the reforms. Although reforms induced women to delay their return to work, the authors do not find that an expansion in maternity leave legislation improves child outcomes. By exploiting unexpected changes in legislation the authors can nicely infer causal effects at the cohort level. We consider our approach complementary to theirs: using individual level instead of cohort data we can evaluate the importance of numerous individual and family characteristics for child outcomes.

A couple of papers use GSOEP data to explain high secondary track attendance, but none analyzes the impact of parental employment. Büchel and Duncan (1998) explore the role of parents' social activities (e.g. socializing with friends, attending cultural events, doing volunteer work), Francesconi, Jenkins and Siedler (2006) the impact of growing up in a family headed by a single mother and Tamm (2007) investigates the effect of parental income on high secondary track attendance.

The structure of the paper is as follows: section 2 offers basic information on the German school system, section 3 provides a brief overview on the GSOEP data set. Economic framework and estimation methods are discussed in section 4. In section 5, we present results from level and sibling difference regressions as well as non-parametric Kernel density estimates. Section 6 concludes.

time on explanatory variables that their original data and the second data set have in common and then apply coefficients to their original data set to construct estimates of child care time.

2 Institutional background: the German school system

In Germany, all students jointly go to elementary school for at least four years. After elementary school, usually at age 10, students proceed to secondary education. The secondary education system is organized in three main tracks: Least academic and most vocational general secondary school track (Hauptschule, grades 5 to 9) provides basic secondary education and prepares for an apprenticeship in a blue collar job. Intermediate secondary school track (Realschule, grades 5 to 10 or 11) is usually followed by an apprenticeship in a white collar job. Only students of the most academic track, the high secondary school track (Gymnasium, grades 5 to 12 or 13), obtain a final degree that provides access to university.

Education is regulated by the states (Bundesländer). In all states, track choice after elementary school is influenced by a recommendation of elementary school teachers that is mainly based on performance and the decision of parents. To which extent parents can influence their child's school track differs substantially across states. In some states the tracking decision is delayed from fourth to sixth grade, and all students jointly go to Förderschule in fifth and sixth grade. Furthermore, some states have a comprehensive school type (Gesamtschule) that comprises all three tracks. All states have schools for children with special needs due to physical or mental disabilities (Sonderschule). Finally, there are very few so-called Waldorf schools that are private and follow a special pedagogy. Still, in our data, about 88 % of students are part of the standard three track system: 20 % attend general secondary school track, 34 % intermediate and high secondary school track each. In all states, secondary education is compulsory up to grade 9 and provided free of charge.

Changing secondary school track after initial choice is possible, but relatively rare. Using GSOEP data on West Germans born between 1970 and 1984, Tamm (2007) compares secondary school tracks attended at age 14 with the highest secondary school degree obtained at age 21. He finds that between 60 % and 70 % of students obtain the degree of the secondary school track they attended at age 14. There is some upward mobility: 21 % (5 %) of those attending intermediate (general) secondary school at age 14 manage to obtain a degree which provides complete or restricted access to university. In each school track roughly 10 % drop out without any degree. Schnepf (2002) provides further evidence on low rates of track changing.

The dependent variable of our analysis is secondary school track, more precisely whether a child attends high secondary track or does not. Secondary school track is an important determinant of labor market outcomes later in life. Using GSOEP data, Dustmann (2004) shows that having successfully attended the high (intermediate) instead of the general secondary school track increases the wage at labor market entry by 29.3 % for men and 37.7 % for women (15.9 % for men and 26.7 % for women, respectively). This holds true even when controlling for post-secondary education that is strongly influenced by secondary school track. The wage premium increases to far more than 50 % for a high instead of general secondary education degree when post-secondary education is not controlled for.

3 Data

We use data from the German Socioeconomic Panel (GSOEP). The GSOEP is a representative panel study of German households that covers the years 1984 until 2006. In addition to household level information, individual information is available. Data cover a wide range of topics such as individual attitudes and health status, job characteristics, unemployment and income, family characteristics and living conditions. For children up to age 15, personal information is provided by the head of the household. We use subsamples A to D, i.e. data on households living in East and West Germany[3] irrespective of their nationality. Haisken-DeNew and Frick (2003) provide a detailed description of the GSOEP.

Our dependent variable is binary and indicates whether a child attends high secondary school track or does not, i.e. attends general or intermediate secondary education. Hence, it focuses on whether children will obtain access to university after finishing school or do an apprenticeship as both general and intermediate secondary track students usually do.[4] We use the latest available information on track choice to minimize inaccurateness caused by later changing of tracks. Children attending other types of schools (such as Gesamtschule, Förderschule, Waldorfschule

[3]Sampling of East German households started in 1990.

[4]Reducing the three track system to a binary dependent variable makes our results better comparable to those of the related literature, see for example Puhani and Weber (2007), Büchel and Duncan (1998) and Francesconi, Jenskin and Siedler (2006). Furthermore, results of a model with a binary dependent variable are easier to interpret than those of an ordered logit model. While it would in principle be possible to estimate an ordered logit model on sibling differences, this would require the additional assumption that the difference between general and intermediate secondary track is the same as the difference between intermediate and high secondary track.

or Sonderschule) are excluded from our analysis.

Parents' time inputs are the primary variables of interest. We use three alternative variables to check the sensitivity of our results: (i) weekly hours worked, (ii) the number of years in which parents have a full-time, part-time or no job, and (iii) hours spent on child care on an average weekday. While average hours spent on child care is the most direct measure, it is also the most subjective one. Some parents claim to devote 24 hours per day to child care, others, who also stay at home, state much lower numbers. In contrast, type of employment and hours worked are more objective measures. They do not capture time inputs directly, but are strongly negatively correlated with hours spent on child care: for fathers, the correlation coefficient ρ between hours worked and time spent on child care is -0.34 and significantly different from zero ($p < 0.001$), for mothers, $\rho = -0.32$ with $p < 0.001$.[5] For the largest part of our analysis we use averages of one of these three variables over a child's first three years. There are two reasons for focusing on the first three years. First, this is the period that is most debated in public - for example, up to now public child care facilities in (West) Germany have nearly exclusively been available for children from age three onwards. Second, as will become clear in the next paragraph, our identifying assumption is that parents do not know their child's ability as long as their child is sufficiently small and, thus, cannot condition their employment decision on their child's ability. This assumption is more plausible the younger a child is.

A list of all explanatory variables, their means and sample sizes is displayed in Table 2. As robustness checks, Tables 8 and 9 in the appendix present the complete results of our main specification when using data on non-foreign West Germans (subsample A) only (Table 8, column (2)), when using school track information at age 14 (instead of latest available information) as dependent variable (Table 8, column (3)) or when using parents' year-specific time use and employment information (Table 9). For all robustness checks, magnitudes of coefficients and significance levels do not change substantially compared to the baseline specification.

[5] Similarly, the correlation coefficients between working full-time and average hours spent on child care per day are $\rho = -0.32$ for fathers and $\rho = -0.28$ for mothers, both with $p < 0.001$.

Table 2: Summary statistics

variable	general sample	siblings sample
information on the child		
attends high secondary school track	0.356	0.329
male	0.500	0.507
year of birth	1989.614	1989.835
firstborn child	0.484	0.402
born from January till June	0.508	0.513
information on the household		
total monthly net equivalent income*,**	0.917	0.924
non-German household	0.238	0.238
information on the mother		
age at birth <=21	0.117	0.127
age at birth 22-35	0.829	0.838
age at birth >36	0.053	0.035
years of education	11.486	11.549
weekly hours worked*	7.743	5.196
time spent on child care per weekday*	8.839	9.454
not working (number of years)	2.189	2.338
part-time job (number of years)	0.477	0.451
full-time job (number of years)	0.334	0.211
information on the father		
age at birth <=21	0.028	0.025
age at birth 22-35	0.829	0.864
age at birth >36	0.143	0.111
years of education	11.949	11.993
weekly hours worked*	40.363	40.182
time spent on child care per weekday*	2.235	2.288
not working (number of years)	0.196	0.198
part-time job (number of years)	0.028	0.048
full-time job (number of years)	2.775	2.754
N***	1047	550

* average at ages 0-3 of child

** in 1000 Euros

*** deviant number of observations for time spent on child care (N=1032 and 537) and for type of employment (N=962 and 521)

4 Economic framework, identification and estimation

Why should parental employment affect children's educational attainments? The very stylized and simplified static framework underlying our empirical analysis assumes that children's educational attainment s_i is a function of parents' time inputs, t_i, and goods and services inputs, x_i, and the child's ability, μ_i: $s_i = f(t_i, x_i, \mu_i)$ where all three first partial derivatives are positive. Both time and good inputs are influenced by parents' employment decisions. On the one hand, we expect parents who work to spend less time with their children (e.g. to play with them or to educate them) which results in a negative "time effect". On the other hand, we expect a positive "input effect". The more parents work the higher is the family income. Due to the income effect normal good inputs such as the number of books and toys at home or extra lessons in the afternoon will increase (if parents are altruistic at least to some degree). In our regressions, we will use family income as (the best available) proxy for goods and services inputs.

Our framework is most closely related to Leibowitz (1974) who assumes that family income has an additional direct impact on the schooling level. A similar relationship can also be derived from a production function framework that draws an analogy between the knowledge acquisition process of an individual and the production process in a firm (see, for example, Todd and Wolpin, 2003). The theory of family behavior (see Becker and Tomes, 1986, 1979 and Solon, 1999 for a simplified version) assumes that parents' intertemporal utility depends on their own consumption and on children's outcomes that are increasing in monetary investments in children. Consequently, parents invest part of their earnings in their children to maximize their own utility subject to a budget constraint. This gives the input effect. The time effect could be obtained by adding a time constraint and time investments to the model.

To begin with we estimate the following logistic regression model for a child i from family j:

$$(1)\ \Pr(high_{ij} = 1|\underline{X_{ij}}, \underline{X_j}) = F(\beta_0 + \underline{\beta_1}\underline{X_{ij}} + \underline{\beta_2}\underline{X_j})$$

where $F(z) = \frac{\exp(z)}{1+\exp(z)}$ is the standard logistic distribution.

$High_{ij}$ is a binary variable that equals one if a child attends or has already finished high secondary track and zero otherwise. $\underline{X_{ij}}$ is a vector of characteristics that differ for different children of one family. It contains (i) child characteristics,

namely a child's year of birth (normalized by subtracting 1984, the first year observed in our data) and binary indicators of a child's sex, whether a child is the firstborn child and whether a child is born between January and June[6], (ii) total net equivalent income of the household averaged over the years 0-3 of the child and (iii) time varying parent characteristics: separate indicators for whether father and mother were younger than 22 or older than 36 when the child was born as well as information on mother's and father's employment or time spent on child care at ages 0-3 of child i.[7] $\underline{X_j}$ is a vector of characteristics that are shared by different siblings of one family j. It encompasses (i) household characteristics, here whether the household is classified as foreign (subsamples B and D in the GSOEP data) and (ii) time invariant characteristics of parents (father's and mother's total years of education measured as schooling plus apprenticeship plus university studies) and (iii) a vector of state dummies. β_0 is a constant and $\underline{\beta_k}, k = 1, 2$ are vectors of unknown parameters.

The coefficients of interest are those of parental employment or time spent on child care. Since we control for household income, they measure the time effect of parents' employment on a child's track choice.

To identify the true underlying coefficients we need to address potential endogeneity problems due to omitted variables. First, since a child's ability is unobserved and coefficients of explanatory variables that are correlated with ability may be inconsistently estimated. In particular, this might be the case for the effect of parental employment if parents condition their employment on their child's ability (reverse causality). To give an extreme example, parents with disabled children might not work at all. We exclude children attending Sonderschule, i.e. disabled children or children with very low ability from our analysis. Apart from these extreme cases, our identification strategy assumes that parents do not know their child's ability as long as their child is sufficiently young and, thus, cannot condition their employment decision on their child's ability. The idea is that information revelation takes time

[6] In Germany, children born between January and June (July and December) usually enter school in autumn of the year in which they become six (seven) years old. Puhani and Weber (2007) show that children who enter school at an older age because they are born between July and December perform better at school and have a higher probability of attending high secondary track.

[7] Averages over household income and time input information are taken over the years in which the information is available, i.e. for some observations we just observe information in one or two out of three years. For both household income and time input variables results are robust if we use only those observations for which information is available for all three years.

and the amount of feedback increases only as a child grows.[8] To make this argument plausible we exclusively use parents' labor market participation when children are aged 0-3. As an additional safeguard, we include parents' education as a proxy variable for their child's ability. Here, we exploit that parents' education is correlated with their own ability which in turn is partially inherited by their children.

Second, selection of parents in the labor market is a potential problem. Imagine that parents with unobserved characteristics u_j that are either especially supportive or detrimental to raising children systematically decide (not) to work. If this were the case, the coefficients of parents' employment would not only capture the time effect that we would like to measure, but also the effect of parents' unobserved characteristics on the child's educational attainment. To control for all time-invariant unobserved parent characteristics we estimate a model on sibling differences (compare Ermisch and Francesconi, 2002 for a similar application). Identification rests on the assumption that parents' relevant unobserved characteristics summarized in u_j (e.g. the quality of parent-child interactions) do not vary for different siblings.

To estimate a model on sibling differences we drop all observations on children without siblings from the sample. We sort all children of a family (who are born between 1984 and 1996 and for whom complete information is available) by age and build pairs of siblings. A pair consists of two adjacent siblings such that we get $n-1$ differences for n siblings. To be able to interpret our results in terms of probability of high secondary track attendance we need to estimate a binary model. For this purpose we construct sorted differences: first, we subtract values of all variables that belong to the older sibling from the values of the corresponding variables of the younger sibling. This results in a new differenced dependent variable that takes values -1, 0 or 1. Second, to be back in a binary framework we multiply all (dependent and explanatory) differenced variables of a sibling pair with -1 if the differenced dependent variable is originally -1 ("sorting"). Equation (2) illustrates this procedure for the oldest two siblings of a family in a linear probability model:

$$(2)\ \Pr\{\mathbf{1}(high_{2j} - high_{1j}) = 1 | \mathbf{1}(\underline{X_{2j}} - \underline{X_{1j}})\} = \widetilde{\beta_0} + \widetilde{\underline{\beta_1}}(\mathbf{1}(\underline{X_{2j}} - \underline{X_{1j}}))$$
$$\text{with } \mathbf{1} = \begin{cases} 1 & if \quad (high_{2j} - high_{1j}) \in \{0, 1\} \\ -1 & if \quad (high_{2j} - high_{1j}) = -1 \end{cases}.$$

The numbers 1 and 2 index the first and second born sibling respectively. By construction, all explanatory variables that have the same value for siblings, i.e. u_j

[8] Our argument is similar to Rosenzweig and Wolpin (1995) who argue that parents update their beliefs about their children's endowments as time passes.

and $\underline{X_j}$, are dropped in a sibling difference estimation. Differences in parents' age at birth are collinear with the difference in years of birth of two siblings and are also dropped. While the original constant term β_0 disappears due to differencing, we include a new constant term $\widetilde{\beta_0}$ to account for the effect of being sorted first (compare Ermisch and Francesconi, 2002 and Ashenfelter and Rouse, 1998). The new constant term $\widetilde{\beta_0}$ arises from differencing a dummy variable that is equal to one if a sibling is sorted first and zero otherwise. $\widetilde{\beta_0}$ captures that due to sorting the sibling sorted first in the difference has a higher probability of attending high secondary track. The interpretation of the sign, but not the level of the coefficients in the difference model is the same as in the level model. For example, imagine that we had estimated a significantly positive coefficient β_k for the effect of the difference in maternal weekly hours worked in a linear probability model. Then β_k would imply that having a mother who works one hour more when sibling 2 (sorted first) is small than when sibling 1 is small increases the probability that sibling 2 attends high secondary track while sibling 1 does not by $\beta_k \times 100$ percentage points. Hence, a positive (negative) sign still stands for a positive (negative) effect.

In contrast to Ermisch and Francesconi (2002), we estimate a linear probability model on the differences and not a logit (or probit) model. The reason is that with a logit or probit model the assumption that the error term has a standard logistic or standard normal distribution will either be true for the original level or the difference model.[9] The main disadvantage of a linear probability model is that it may predict probabilities larger than unity or smaller than zero for extreme values of explanatory variables. Since we are interested in the average marginal effects this is not a major problem. Furthermore, we will check how close the estimates of the average marginal effects in a linear probability model are to those obtained in a probit model.

A more commonly used alternative to a sibling difference model is a household fixed effect model (conditional logit model). In our application, a conditional logit model uses only those observations on sibling pairs in which one sibling attends high secondary track and the other one does not. It estimates coefficients by comparing sibling pairs in which the older sibling attends high secondary track to sibling pairs in which the younger sibling attends high secondary track. Thus, identification of the effect of parental employment on children's educational attainment stems from ob-

[9] The variance of the difference of two random variables is equal to the sum of the two variances minus the covariance. Thus, the variance of the difference of two random variables with a standard logistic (normal) distribution only equals the variance of the standard logistic (normal) distribution if the covariance incidentally coincides with the standardized variance.

servations in which both educational outcome and parental employment differ across siblings. In contrast, our sibling difference model also uses observations from families in which all children go to the same secondary track and estimates coefficients by comparing siblings pairs in which both siblings do or do not attend high secondary track to sibling pairs in which only one sibling attends high secondary track. Consequently, the identification of the effect of parental employment on children's educational attainment stems from all observations in which parental employment differs across siblings. We prefer estimating a sibling difference model to estimating a household fixed effect model because the former uses more observations which allows estimating coefficients more precisely.

5 Results

5.1 Estimation on levels

We will first present results from a logit estimation (Table 3) that does not address endogeneity problems caused by unobserved parent characteristics. The results are still a useful benchmark for comparison with other studies that use similar specifications. Additionally, results provide some information on the coefficients of variables that are constant for siblings and thus will drop out when estimating a sibling difference model.

While the coefficient of mother's average hours worked is not significant, the coefficient of father's average hours worked is weakly significant (p=0.080) and positive. Setting all control variables to their mean the predicted marginal effect if the father would work one hour more per week in every year is a 0.5 % increase in the probability that his child attends high secondary track. Furthermore, male children are predicted to attend high secondary track less often than female children. Firstborn children are more likely, children born between January and June are marginally less likely to attend high secondary track. In Puhani and Weber (2007) children who enter school at an older age because they are born between July and December have a 12 % higher probability of attending high secondary track. If control variables are set to their mean, our results predict a 9 % higher probability of attending high secondary track. Having a father who is at most 21 in the year of birth is predicted to have an adverse effect on the child's educational attainment, having a relatively old mother seems to be supportive. Both coefficients are likely to suffer from endogeneity problems and hence might reflect unobserved parents' characteristics that

Table 3: Base specification: logit estimation on levels

binary dependent variable: child attends high secondary track

explanatory variables	coefficient	p-value
mother's weekly hours worked*	0.000	0.960
father's weekly hours worked*	0.019	0.080
male	-0.639	0.008
born before July	-0.380	0.095
firstborn child	0.635	0.003
year of birth - 1984	-0.006	0.889
age of mother at birth <=21	0.598	0.133
age of mother at birth >36	0.987	0.034
age of father at birth <=21	-3.822	0.000
age of father at birth >36	0.259	0.487
mother's total years of education	0.420	0.000
father's total years of education	0.348	0.000
household income*,**	0.086	0.927
(household income)^2 *,**	-0.080	0.785
non-German household	0.180	0.690
constant	-11.136	0.000
state dummies	yes	
N	1047	
Pseudo R^2	0.393	

* average at ages 0-3 of child

** total monthly net household equivalent income in 1000 Euros

comment: robust, clustered standard errors that allow observations to be correlated within a family

are correlated with the included age intervals. The coefficients of parents' total years of education are highly significant and positive. We use state dummies to control for state specific differences in shares of students attending high secondary track.[10]

Table 4 presents the coefficients of parental time inputs. The upper part uses hours worked per week averaged over the ages 0-3 of a child as measure of time inputs, the middle part hours spent on child care on a typical weekday averaged over the ages 0-3 of a child. The lower part presents coefficients of type of employment, e.g. variables that indicate how many out of a child's first three years parents did work full-time, part-time or not at all. The omitted categories are the most common ones: working full time for fathers and not working at all for mothers. The table contains the coefficients from three different specifications. Specification A uses observations from families in which both parents are present (in our data) and information on both parents' time inputs, age and education is completely available. As most studies on the relationship between parental employment and children's educational attainment specification B does not include information on father's age, education and time inputs as explanatory variables. Hence, it additionally includes observations from families with single mothers and families with present fathers on whom information is incomplete which raises the number of observations by about 20 %. We add a dummy variable for absent fathers that turns out not to be significant. Specification C adds a dummy variable that takes a value of one if the father is present and zero otherwise and reports coefficients of father's age, education and time inputs when age, education and time inputs are interacted with this dummy. Otherwise explanatory variables in Table 4 are the same as in Table 3.

In the eight additional specifications, all coefficients of parental time inputs are not significant with one exception: the coefficient of mother's full time employment is negative and significant (p=0.022) in specification B.

[10] Results reported in Table 3 are very robust to using mother's and father's age and age squared instead of age intervals, including year dummies instead of imposing a linear time trend or to including dummies for the number of siblings which we do not do because it reduces the number of observations by about 10 %. Furthermore, using an ordered logit specification with a dependent variable that takes the values 1 to 3 for general, intermediate and high secondary track attendance produces estimates very similar to those reported in Table 3.

Table 4: Further specifications: logit estimation on levels

binary dependent variable: child attends high secondary track

coefficients of average weekly hours worked

specification	(A)	(B)	(C)
mother	0.000	-0.008	0.006
	(0.960)	(0.327)	(0.487)
father	0.019	-	0.012
	(0.080)	-	(0.297)
N	1047	1280	1214
Pseudo R^2	0.393	0.305	0.354

coefficients of type of employment in years

specification	(A)	(B)	(C)
full time mother	-0.189	-0.253	-0.198
	(0.210)	(0.022)	(0.158)
part time mother	0.044	-0.044	0.088
	(0.795)	(0.739)	(0.568)
part time father	-0.422	-	-0.493
	(0.530)	-	(0.407)
non-working father	0.200	-	0.239
	(0.352)	-	(0.276)
N	962	1195	1118
Pseudo R^2	0.388	0.335	0.372

coefficients of average hours spent on child care per weekday

specification	(A)	(B)	(C)
mother	0.050	0.031	0.005
	(0.156)	(0.300)	(0.867)
father	-0.110	-	-0.096
	(0.092)	-	(0.157)
N	1032	1262	1199
Pseudo R^2	0.387	0.302	0.348

comments: robust, clustered standard errors that allow observations to be correlated within a family; p-values are reported in brackets
(A): uses only observations with complete information on both parents
(B): uses all observations with complete information on mother
(C): estimates coefficients of father's age, education and time inputs conditional on father being present

5.2 Estimation on sibling differences

5.2.1 The sample

The siblings sample contains data on 550 siblings from 249 families. Table 2 compares means in the general and the siblings sample. The siblings sample is largely representative for the general sample. Differences in means usually occur only in the second position after the decimal point. Of course, the sibling sample contains fewer firstborn children (40 % instead of 48 %). On average, mothers in the siblings sample work 2.5 hours less per week and spend 0.6 additional hours per day on child care. Father's employment is very similar in both samples.

5.2.2 Kernel density estimates

To get a first impression whether differences in parental employment could be driving differences in siblings' educational attainment we estimate non-parametric Kernel densities. The solid line depicts sibling pairs who either both attend high secondary track or both do not. The dashed line stands for sibling pairs in which one sibling attends high secondary track, but the other one does not. Again, the sibling attending high secondary track is sorted first in the difference. Figure 1 (Figure 2) displays Kernel density estimates of the distributions of differences in average hours worked by mothers (fathers) when children were aged 0-3 for these two kinds of sibling pairs. If having a mother or father with longer working hours would reduce the attendance of high secondary track we would expect the dashed line to be first order stochastically dominated by the solid line.

Eyeballing suggests that estimated densities are very similar. Non-parametric, two-sided Mann-Whitney and Kolmogorov-Smirnov tests on the original (not the displayed estimated) distributions confirm that distributions do not differ significantly: for mothers $p_{MW} = 0.394$ and $p_{KS} = 0.824$ and for fathers $p_{MW} = 0.740$ and $p_{KS} = 0.404$. At first sight, differences in parental employment patterns do not seem to be a driving force behind different levels of educational attainments.[11]

[11] Appendix 7.1 displays estimated Kernel densities for the average time parents spent on child care.

Figure 1: Kernel density estimates, mother's hours worked

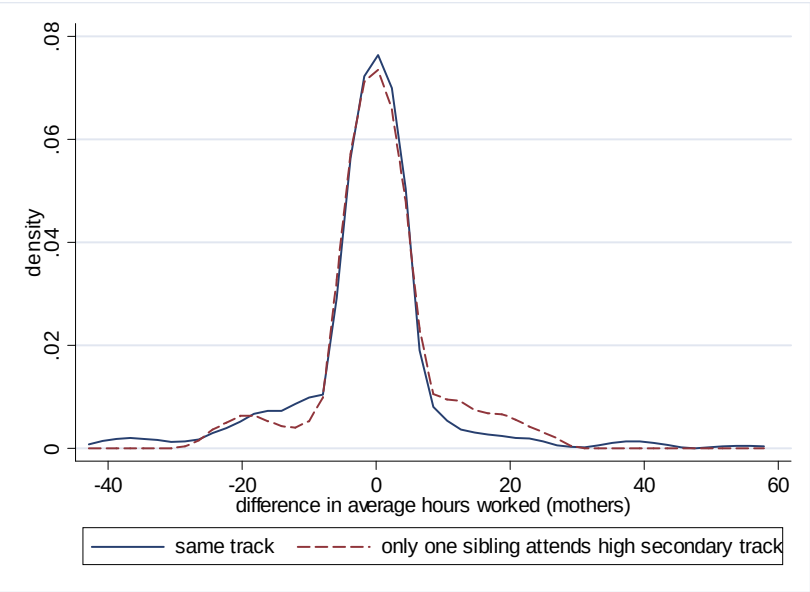
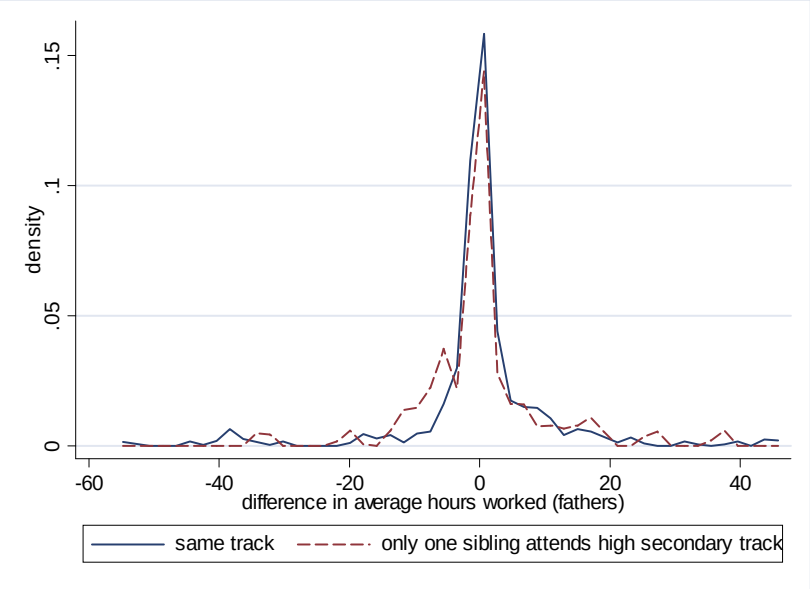


Figure 2: Kernel density estimates, father's hours worked



5.2.3 Multivariate analysis

To control for differences between siblings apart from parental employment we estimate a linear probability model on sibling differences to explain different educational outcomes. The estimation requires sufficient variation in both dependent and explanatory variables. In all specifications, we have about 20 % of sibling pairs in which just one sibling attends high secondary track. Table 5 and Figures 1-4 document substantial variation in mother's and father's average hours worked as well as in time spent on child care. By construction variation is largest in working hours that are measured per week, followed by hours spent on child care measured at a daily level. Variation is smallest for type of employment that is measured in years such that differences can at most range from -3 to 3. For this reason we will provide estimates on sibling differences only for the former two measures of parental time inputs (in contrast to Ermisch and Francesconi (2002) who use the difference in type of employment).

Table 5: Variation in key explanatory variables

differenced variable	mean	standard deviation	zeros (%)	min	max	N
mother's hours worked	-.822	9.960	57.48	-40	55	301
father's hours worked	.106	10.732	17.28	-54	45	301
mother's hours spent on child care	.525	3.497	16.61	-10	19	295
father's hours spent on child care	-.167	2.008	17.63	-9	10	295
mother's full time employment	-.076	.601	86.59	-3	3	276
mother's part time employment	.007	.772	70.65	-3	3	276
non-working mother	.069	.906	65.22	-3	3	276
father's full time employment	.025	.581	87.68	-3	3	276
father's part time employment	-.007	.256	97.46	-3	2	276
non-working father	-.007	.490	90.58	-3	3	276

For our baseline specification, Table 6 compares the marginal effects predicted by an estimation on sibling differences using the linear probability or the probit model. The marginal effects - especially the significant ones - are very similar. Thus, the usual drawbacks of estimating a linear probability model are not a major concern in our application and for the reasons outlined in section 4 we prefer the linear probability model.

Table 6: Linear probability and probit model on sibling differences

dependent variable: sibling difference in high secondary track attendance

model	linear probability			probit	
differenced variables	coefficient	p-value	95 % confidence interval	marginal effects*	p-value
mother's hours worked**	0.005	0.105	[-0.001, 0.012]	0.005	0.073
father's hours worked**	0.004	0.529	[-0.008, 0.015]	0.002	0.589
male	0.009	0.715	[-0.040, 0.058]	0.016	0.488
born before July	-0.097	0.032	[-0.185, -0.009]	-0.093	0.008
firstborn child	0.151	0.003	[0.050, 0.251]	0.089	0.015
year of birth - 1984	-0.042	0.005	[-0.071, -0.012]	-0.029	0.030
household income***	0.019	0.961	[-0.740, 0.777]	0.032	0.928
(household income)^2 ***	0.026	0.837	[-0.221, 0.273]	0.003	0.977
constant	0.317	0.000	[0.228, 0.406]	-0.063	0.000
N	301			301	
R^2	0.240			0.251	

* all other explanatory variables are evaluated at their mean

** average per week at ages 0-3 of child

*** total monthly net household equivalent income in 1000 Euros, average at ages 0-3 of child

comment: robust, clustered standard errors that allow observations to be correlated within a family

Results in Tables 6 and 7 show that differences in father's employment do not contribute significantly to explaining differences in educational attainment. The coefficient of differences in mother's employment is positive and just not significant (in Table 7, p=0.105 in specification (A) and (C) and p=0.085 in specification (B)). The precision of our baseline estimate (specification (A)) implies that we can statistically rule out that having a mother who works one hour more per week (when the sibling sorted first was young than when the second sibling was young) lowers the probability that the sibling who is sorted first attends high secondary track (while the second sibling does not) by more than 0.1 percentage points. Similarly, the alternative specifications in Table 7 show that average time spent on child care does not influence attendance of high secondary track significantly.[12] As in the level

[12]Results in Table 7 are robust to adding a squared term for mother's and father's time inputs.

estimation our sibling difference estimates in Table 6 predict children who are born between January and June to be less likely and firstborn children to be more likely to attend high secondary track, but the advantage of firstborn children decreases with each year they are apart from the second born sibling. In contrast to the level estimation, the sex of a child is no longer significant.

Table 7: Further specifications: linear probability model on sibling differences

dep. var.: sibling difference in high secondary track attendance

coefficients of difference in hours worked, average per week

specification	(A)	(B)	(C)
mother	0.005	0.006	0.005
	(0.105)	(0.085)	(0.105)
father	0.004	-	0.003
	(0.529)	-	(0.582)
N	301	372	345
R^2	0.240	0.307	0.321

coefficients of difference in time spent on child care,

average per weekday

specification	(A)	(B)	(C)
mother	-0.003	-0.003	-0.002
	(0.747)	(0.688)	(0.772)
father	-0.016	-	-0.012
	(0.361)	-	(0.478)
N	295	361	338
R^2	0.239	0.293	0.314

comments: robust, clustered standard errors that allow observations to be correlated within a family; p-values are reported in brackets
(A): uses only observations with complete information on both parents
(B): uses all observations with complete information on mother
(C): estimates coefficients of father's age, education and time inputs conditional on father being present

The size of coefficients in the level and the difference estimation is not directly comparable since they refer to characteristics measured in levels or differences between siblings, respectively. Still, the interpretation of coefficients' signs (i.e. whether we observe a negative or positive effect) is comparable. The signifi-

cance levels and, thus, implications from the level and difference estimation differ markedly. Table 2 documents that differences are not caused by different sample characteristics. This suggests that controlling for unobserved parent characteristics affects results and should become the standard in the literature on the effects of parental employment on children's educational attainment.

Furthermore, mother's and father's time inputs do not seem to influence children's educational outcomes in different ways: we can reject the hypothesis that the coefficients of mother's and father's employment and time spent on child care differ significantly (F-tests for specifications (A) in Table 7 yield F=0.09, p=0.763 for hours worked and F=0.33 and p=0.565 for time spent on child care). Table 8, Columns (4) and (5) in Appendix **??** display estimates for parents' joint working hours and joint time spent on child care. Since the coefficients of joint time inputs are not significant, they confirm our previous results.

Additionally, we check whether the effect of parental employment differs at different ages of the child, in our case at age 1, 2 and 3 (see Appendix **??**, Table 9). Some studies that focus on children's short term outcomes have found that maternal employment during the first year of a child is especially detrimental. For example, Ruhm's (2000) and Waldfogel, Han and Brooks-Gunn's (2002) results imply that maternal employment during the first year of a child reduces math, reading and verbal achievement test scores at the ages 3-8 substantially. Our results on long-term outcomes are not consistent with this finding. In contrast, the coefficient of maternal working hours during the first year is marginally significant (p=0.082) and positive. F-tests for equality of coefficients document that each parent's coefficients do not differ across the ages 0-3 of a child. This justifies our approach of averaging time input information over the first three years.[13]

6 Concluding remarks

This paper has analyzed whether parental employment affects children's educational attainment. We have explicitly addressed potential endogeneity problems: to control for unobserved parent characteristics we have used estimates on sibling differences.

[13]For the three coefficients of mother's (father's) hours worked F=1.02 and p=0.364 (F=0.41 and p=0.666), for the three coefficients of mother's (father's) time spent on child care F=0.54 and p=0.586 (F=0.06 and p=0.939). Furthermore, in both specifications all six parents' coefficients do not differ significantly (F=0.58 and p=0.713 for hours worked and F=0.42 and p=0.832 for time spent on child care, respectively).

To avoid inconsistent estimates due to reverse causality we have dropped disabled children from the analysis, have focused exclusively on parental employment when children are young (aged 0-3) such that signals about ability are still scarce and have included parent's education as a proxy variable.

Our measures of parental time inputs exclusively capture quantity, not quality - though quality is controlled for in the sibling difference estimates to the extent quality of parent-child interactions does not differ for different siblings. Due to lack of data, we have not controlled for non-parental time and good inputs and have ignored potentially important differences between different kinds of non-parental child care (such as attendance of Kindergarten, child care by relatives or nannies). These are important issues left for future research. Still, it is often argued that parental employment patterns per se shape a child's environment and outcomes. This is what we have tested for.

In sum, our results do not support worries that parental employment is detrimental for children's educational attainment. The core of our analysis are the estimates on sibling differences that use average weekly working hours when the child is aged 0-3 to measure parental time inputs: given their precision, we can statistically rule out that having a mother who works one hour more per week lowers the probability of high secondary track attendance by more than 0.1 percentage points, an economically negligible number. Actually, all coefficients of maternal employment are positive but not significant at conventional significance levels (though at an 9 to 11 % level). The corresponding coefficients of paternal employment and estimates using parental time spent on child care instead of working hours are not significant. Taken together, our results imply that it is not parental employment or quantity of parent-child interactions that are decisive for children's educational attainments, but, for example, birth order within a family, age relative to classmates or parental characteristics.

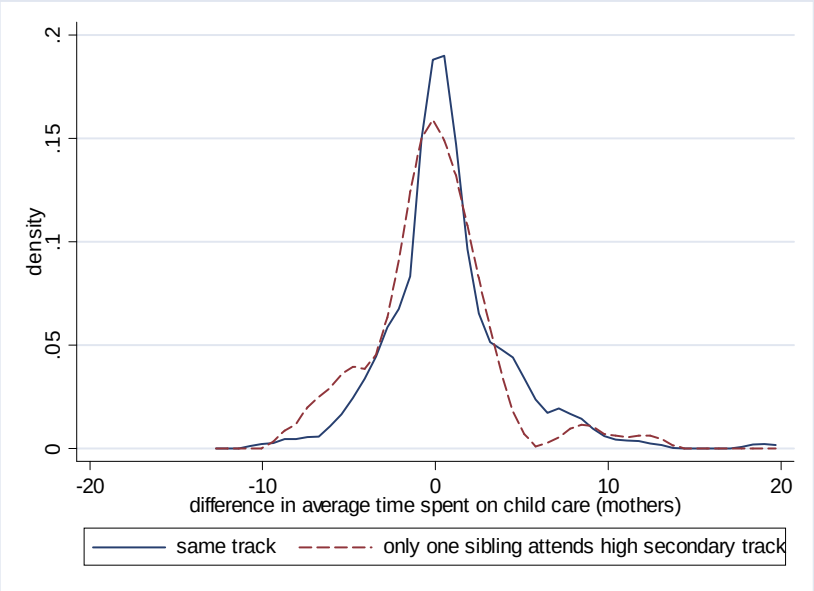
With respect to the current debate about the expansion of day care facilities in Germany our results do clearly not support worries that a more comprehensive child care infrastructure will hurt children's future prospects by raising maternal employment. Of course, our estimates are based on data from the past. To some extent, the current reforms will lead to changes in the institutional environment and perhaps also society's attitudes towards working mothers that may affect the interplay between parental employment and child outcomes.

7 Appendix

7.1 Kernel density estimates for time spent on child care

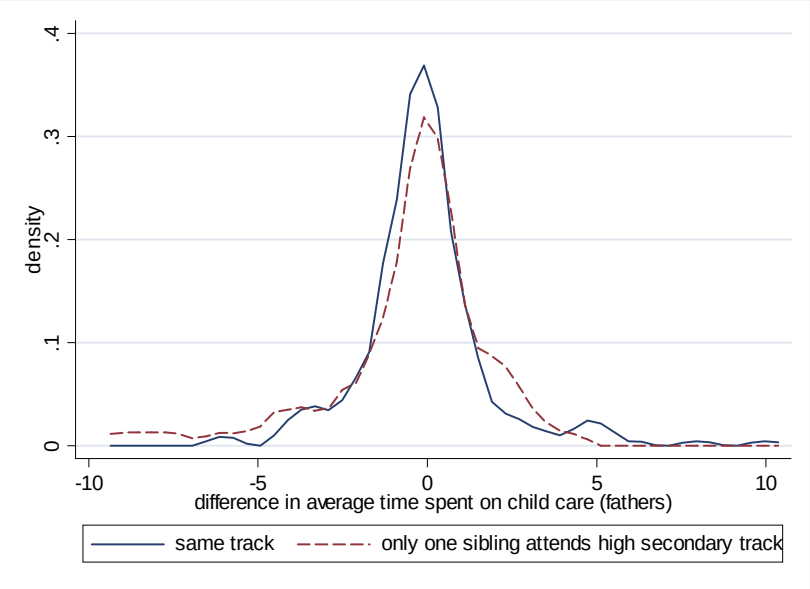
Figure 3 (Figure 4) displays Kernel density estimates of the distributions of differences in average hours spent on child care by mothers (fathers). The dashed line depicts sibling pairs in which one sibling attends high secondary track and the other one does not. The solid line marks siblings who either both attend high secondary track or both do not.

Figure 3: Kernel density estimates, mother's time spent on child care



Non-parametric, two-sided Mann-Whitney and Kolmogorov-Smirnov tests on the original distributions confirm that distributions do not differ significantly for fathers, $p_{MW} = 0.953$ and $p_{KS} = 0.884$. In contrast, distributions of mother's time spent on child care differ marginally: $p_{MW} = 0.076$ and $p_{KS} = 0.099$. Regression results in Table 7 show that differences in the time that mothers spent on child care cannot explain different educational attainment of siblings when other explanatory variables are controlled for.

Figure 4: Kernel density estimates, father's time spent on child care



7.2 Robustness checks

Table 8: Robustness checks I

dependent variable: sibling difference in high secondary track attendance

specification	(1)	(2)	(3)	(4)	(5)
sample	general	West Germ.	general	general	general
dependent variable	latest obs.	latest obs.	at age 14	latest obs.	latest obs.
differenced variables					
mother's weekly hours	0.005	0.006	0.004	-	-
worked*	(0.105)	(0.113)	(0.268)	-	-
father's weekly hours	0.004	0.004	0.001	-	-
worked*	(0.529)	(0.629)	(0.892)	-	-
parents' joint weekly	-	-	-	0.004	-
hours worked*	-	-	-	(0.180)	-
parents' joint hours	-	-	-	-	-0.007
spent on child care*	-	-	-	-	(0.353)
male	0.009	0.006	-0.012	0.009	0.011
	(0.715)	(0.884)	(0.780)	(0.732)	(0.683)
born before July	-0.097	-0.112	-0.127	-0.094	-0.085
	(0.032)	(0.033)	(0.029)	(0.024)	(0.027)
firstborn child	0.151	0.151	0.165	0.148	0.168
	(0.003)	(0.057)	(0.007)	(0.002)	(0.001)
year of birth	-0.042	-0.043	-0.042	-0.043	-0.040
	(0.005)	(0.016)	(0.096)	(0.001)	(0.003)
household income**	0.019	0.074	0.379	0.030	0.174
	(0.961)	(0.867)	(0.585)	(0.935)	(0.715)
(household income)^2 **	0.026	0.006	-0.070	0.023	-0.022
	(0.837)	(0.986)	(0.748)	(0.849)	(0.883)
constant	0.317	0.330	0.374	0.317	0.327
	(0.000)	(0.050)	(0.000)	(0.000)	(0.000)
N	301	213	163	301	295
R^2	0.240	0.251	0.220	0.234	0.236

* average at ages 0-3 of child
** total monthly net equivalent income in 1000 Euros, average at ages 0-3 of child
comments: robust, clustered standard errors that allow observations to be correlated within a family; p-values are reported in brackets

Table 9: Robustness checks II

dependent variable: sibling difference in high secondary track attendance

key explanatory variables differenced variables	difference in weekly hours worked	difference in daily hours spent on child care
mother's time input at age 1	0.004	-0.001
	(0.082)	(0.860)
mother's time input at age 2	0.000	0.007
	(0.970)	(0.270)
mother's time input at age 3	0.000	-0.005
	(0.867)	(0.515)
father's time input at age 1	0.001	-0.002
	(0.754)	(0.891)
father's time input at age 2	0.003	-0.001
	(0.236)	(0.903)
father's time input at age 3	0.001	-0.008
	(0.824)	(0.708)
male	0.002	0.011
	(0.949)	(0.835)
born before July	-0.053	0.015
	(0.184)	(0.741)
firstborn child	0.164	0.150
	(0.010)	(0.018)
year of birth	-0.061	-0.058
	(0.000)	(0.000)
household income*	0.530	-0.147
	(0.234)	(0.807)
(household income)^2 *	-0.099	0.458
	(0.485)	(0.458)
constant	0.380	0.336
	(0.000)	(0.000)
N	219	108
R^2	0.326	0.416

* total monthly net equivalent income in 1000 Euros, average at ages 0-3 of child
comments: robust, clustered standard errors that allow observations to be correlated within a family; p-values are reported in brackets

References

[1] Ashenfelter, O., Rouse, C., 1998. Income, schooling, and ability: evidence from a new sample of identical twins. Quarterly Journal of Economics 113 (1), 253-284.

[2] Becker, G. S., Tomes, N., 1979. An equilibrium theory of the distribution of income and intergenerational mobility. Journal of Political Economy 87 (6), 1153-1189.

[3] Becker, G. S., Tomes, N., 1986. Human capital and the rise and fall of families. Journal of Labor Economics 4 (3) Part 2, 1-39.

[4] Blank, R. M., 2002. Evaluating welfare reform in the United States. Journal of Economic Literature 40 (4), 1105-1166.

[5] Buchel, F., Duncan, G. J., 1998. Do parents activities promote chidren's school attainments? Evidence from the German Socioeconomic Panel. Journal of Marriage and the Family 60 (1), 95-108.

[6] Chao, E. L., Rones, P. L., 2007. Women in the labor force: a databook. U.S. Department of Labor, U.S. Bureau of Labor Statistics, http://www.bls.gov/cps/wlf-databook-2007.pdf.

[7] Dustmann, C., 2004. Parental background, secondary school track choice, and wages. Oxford Economic Papers 56, 209-230.

[8] Dustmann, C., Schönberg, U., 2007. The effect of expansions in maternity leave coverage on children's long-term outcomes. Mimeo.

[9] Ermisch, J., Francesconi, M., 2002. The effect of parents' employment on children's educational attainment. ISER working paper 2002-21.

[10] Ermisch, J., Francesconi, M., 2005. Parental employment and children's welfare. In: Boeri, T., del Boca, D., Pissarides, C. (Eds.). Women at work: an economic perspective. Oxford University Press, Oxford, 154-193.

[11] Francesconi, M., Jenkins, S. P., Siedler, T., 2006. Childhood family structure and schooling outcomes: evidence for Germany. DIW Discussion paper 610.

[12] Graham, J. W., Beller, A. H., Hernandez, P. M., 1994. The effects of child support on educational attainment. In: Garfinkel, I., McLanahan, S. S., Robins, P. K. (Eds.). Child support and child well-being. The Urban Institute Press, Washington, 317-349.

[13] Haisken-deNew, J. P., Frick, J. R. (Eds.), 2003. DTC Desktop Companion to the German Socio-Economic Panel (SOEP). DIW, Berlin.

[14] Haveman, R., Wolfe, B., 1995. The determinants of children's attainments, a review of methods and findings. Journal of Economic Literature 33 (4), 1829-1878.

[15] Hill, M. S., Duncan, G. J., 1987. Parental family income and the socioeconomic attainment of children. Social Science Research 16, 39-73.

[16] Kiernan, K., 1996. Lone motherhood, employment and outcomes for children. International Journal of Law, Policy and the Family 10, 233-249.

[17] Leibowitz, A., 1974. Home investments in children. Journal of Politcal Economy 82 (2) Part 2, 111-131.

[18] O'Brien, M., Jones, D., 1999. Children, parental employment and educational attainment: an English case study. Cambridge Journal of Economics 23, 599-621.

[19] Puhani, P. A., Weber, A. M., 2007. Does the early bird catch the worm? Instrumental variable estimates of early educational effects of age of school entry in Germany. Empirical Economics 32, 359-386.

[20] Rosenzweig, M. R., Wolpin, K. I., 1995. Sisters, siblings, and mothers: the effect of teen-age childbearing on birth outcomes in a dynamic family context. Econometrica 63 (2), 303-326.

[21] Ruhm, C. J., 2004. Parental employment and child cognitive development. Journal of Human Resources 39 (1), 155-192.

[22] Schnepf, S. V., 2002. A sorting hat that never fails? The transition from primary to secondary school in Germany. Innocenti working paper No. 92.

[23] Solon, G., 1999. Intergenerational mobility in the labor market. In: Ashenfelter, O., Card, D. (Eds.). Handbook of Labor Economics, Vol. 3. Elsevier Science, 1761-1800.

[24] Tamm, M., 2007. Does money buy higher schooling? Evidence from secondary school track choice in Germany. ECINEQ Working paper 2007-58.

[25] Todd, P. E., Wolpin, K. I., 2003. On the specification and estimation of the production function for cognitive achievement. Economic Journal 113 (485), 3-33.

[26] Waldfogel, J., Han, W.-J., Brooks-Gunn, J., 2002. The effects of early maternal employment on child cognitive development. Demography 39 (2), 369-392.