First draft May 1998

Revised August 14, 1998

Does Family Background Matter? - Returns to Education and Family Characteristics in Germany

Isabel Gödde and Reinhold Schnabel

University of Mannheim, Germany[1]

SFB 504 Discussion Paper # 98-60

Abstract:

Using data from the GSOEP, we analyze the impact of parents' education („family background") on wages and on returns to education. We show that returns to education are heterogeneous with family background accounting for part of this heterogeneity. While the level of wages is higher for individuals from families with higher education, estimated returns to education are higher for individuals from families with lower education levels. However, a simple regression analysis suffers from serious endogeneity problems. The construction of a sibling sample allows us to perform a family fixed-effect estimation in order to control for unobserved characteristics that are shared by family members. Our results suggest that the conventional estimates overstate the returns to education. Moreover, family background accounts for a large part of the variation in wages in the sibling sample.

---

[1] Financial support from the *Deutsche Forschungsgemeinschaft* and the *Deutscher Akademischer Auslandsdienst* (DAAD) is gratefully acknowledged. We would like to thank the Department of Economics at the University of California, Berkeley, for its hospitality and especially Kenneth Chay for helpful advice. Moreover, we would like to thank Axel Börsch-Supan and the participants of the Brown Bag Seminar at the University of Mannheim for comments.

I. Introduction

In the 1960s substantial changes took place in the German educational system. The reforms aimed at an equalization of education opportunities in all social classes and a decrease in the impact of family background on children's education. One consequence was the abandonment of school tuition fees as well as the attempt to equalize school quality allover the country. An interesting question is now: Does family background still play an important role in the determination of children's education and wages in Germany?

In this paper we try to answer this question by examining the role of family background in the human capital production function, i.e. its influence on wages as well as on returns to education. In the analysis we use a simple quantity measure of education that, however, also captures some important differences in school quality in Germany, as will be explained below. We do not consider the effect of school quality explicitly.

While this area of research has been very active in the United States in recent years, empirical work in Germany has focussed on questions of industry-specific wage structure, wage dispersion, inequality and wage stickiness (Abraham and Houseman 1995, Fitzenberger et al. 1995, Nickell and Bell 1996, Steiner and Wagner 1996). An earlier paper by Licht and Steiner (1991) addresses the problems of individual heterogeneity in the estimation of human capital earnings functions. Licht and Steiner try to control for unobserved heterogeneity using panel data. In order to identify the returns to education they rely on changes in education over time. Unfortunately this procedure is likely to pick up measurement error, and moreover, is treating persons with uncompleted education histories as if the transitory education status were optimal.

In order to reduce the bias in the estimation of returns to education we (i) introduce family background variables as additional controls, (ii) allow for heterogeneous returns to education, and (iii) construct a sibling sample from the German Socio-Economic Panel (GSOEP) and use a fixed effects estimator in order to control for unobserved family-specific heterogeneity. This potentially reduces the bias in the estimation of returns to education.[2]

There are two remaining problems that we do not solve in this paper. First, the fixed-effect estimation does not eliminate idiosyncratic differences among the siblings. This induces an

---

[2] Altonji and Dunn (1996) have done a similar analysis for the United States using PSID and NLS data.

upward bias into our estimates of returns to education. Second, we do not tackle the problem of measurement error in the education variable. This measurement error might lead to an attenuation bias in the estimates that might offset the above-mentioned bias at least partially.[3] In the future we will try to find a solution to the latter problem, while the former problem is probably going to remain unsolved.

Our results can be summarized as follows: In the full sample family background appears to play an important role in the determination of wages. Moreover, returns to education seem to be heterogeneous with family background being an important determinant of this heterogeneity. Persons from less educated families tend to have lower wages, but higher returns. However, the estimates of returns to education are very likely to be biased upwards because of an endogeneity bias. This problem is tackled in the sibling analysis. In the OLS regressions using the sibling sample the effect of family background is weaker than in the full-sample regressions and returns to education appear to be homogeneous. Most strikingly, the fixed-effects estimates of returns to education are much smaller than the ones from the OLS analysis, with the coefficient of education becoming insignificant. Moreover, family background accounts for a large part of the variation in wages.

The paper is organized as follows: In section II, we derive the econometric model that is the basis of our analysis. In the next section, we discuss specification issues and suggest solutions to the endogeneity problem. In section IV the data set, the variables, and the samples are described. Section V proceeds with presenting and interpreting the result. First we estimate the earnings function based on the full sample and control for heterogeneity in the level of wages by using family background variables. Then we also include interactions of education with other variables. Finally, we present estimates of the earnings function based on the sibling sample. Section VI concludes.

[3] However, we believe that measurement error is less of a problem when comparing education across siblings than when comparing education for a given person over time.

II. Empirical Model

The research on returns to education is based on the work by Becker (1967) and Mincer (1974). In the traditional specification, returns to education are estimated as follows:

$$\log(y_i) = a + bS_i + cE_i + dE_i^{\,2} + \varepsilon_i\,, \tag{1}$$

where y_i is average lifetime income, S_i is years of completed education, E_i is experience, and ε_i is a statistical residual. This function that has been introduced by Mincer (1974) is known as the *human capital earnings function*. It has been the basis of practically all research on returns to education. Recently, economists have begun to extend the simple model to capture additional aspects of returns to education. The framework used in this paper is due to Card (1995, 1998)[4]. It differs from the traditional approach in that it allows for heterogeneous returns to education.

The underlying assumption of most models on human capital formation is that agents maximize their life-time utility that depends on the average level of earnings over the lifecycle, denoted by y(S), and the disutility from education, h(S). This is specified as follows:

$$\begin{aligned} &\max_{y,S}\; U(S_i, y_i) = \log(y_i) - h(S_i) \quad s.t.\; y_i = y(S_i) \\ &\Leftrightarrow \max_{S} U(S_i) = \log(y(S_i)) - h(S_i) \end{aligned} \tag{2}$$

Linearity in log earnings implies that the optimal education choice does not depend on factors that raise earnings proportionally for <u>all</u> levels of education. $h(S_i)$ is increasing and convex in S_i.

The first-order condition of this maximization problem sets the marginal benefits of education equal to its marginal costs:

$$\frac{y'(S_i^*)}{y(S_i^*)} = h'(S_i^*) \tag{3}$$

Individual heterogeneity is introduced into this model by allowing marginal benefits and marginal costs to vary across individuals. Marginal benefits are assumed to be linear and decreasing in S_i where the intercept is individual-specific:

---

[4] This section follows closely the presentation in Card (1998).

$$\frac{y'(S_i)}{y(S_i)} = b_i - k_1 S_i \tag{4}$$

By assumption, marginal costs are also linear in S_i and increasing. Heterogeneity enters again through the intercept term:

$$h'(S_i) = r_i + k_2 S_i \tag{5}$$

The optimal level of education derived from the first-order condition (2) is:

$$S_i* = \frac{b_i - r_i}{k_1 + k_2} \tag{6}$$

At the optimum, the marginal return to education of individual i is:

$$\beta_i = b_i - k_1 S_i* = b_i (1 - \frac{k_1}{k_1 + k_2}) + r_i \frac{k_1}{k_1 + k_2} \tag{7}$$

It is easily derived that returns to education will differ across individuals unless one of the following two conditions is satisfied:

1. marginal benefits are constant and equal for all i (i.e. $b_i = \overline{b}$ $\underline{and}$ $k_1 = 0$) or
2. marginal costs are constant and equal for all i (i.e. $r_i = \overline{r}$ $\underline{and}$ $k_2 = 0$).

By integrating equation 4, we can derive an econometric model that is very similar to the basic Mincer equation:

$$\log(y_i) = \alpha_i + b_i S_i - \frac{1}{2} k_1 S_i^{\,2} + \varepsilon_i \tag{8}$$

The important differences to Mincer's model are the individual-specific intercept and slope terms that turn the model into a *random-coefficient model*. The equation can be written in the form of deviations from means in the following way:

$$\log(y_i) = a_0 + \overline{b} S_i - \frac{1}{2} k_1 S_i^{\,2} + a_i + (b_i - \overline{b}) S_i + \varepsilon_i\ , \tag{9}$$

where $a_i = \alpha_i - a_0$ with mean 0.

The basic specification in our empirical work is a modified version of equation 9:

$$\log(y_i) = \alpha + \beta_1 S_i + \beta_2 X_i S_i + \beta_3 S_i^{\,2} + \gamma_1 E_i + \gamma_2 E_i^{\,2} + \delta X_i + \varepsilon_i \tag{10}$$

where S_i is years of education, E_i is working experience, and X_i are other controls that will be described below. The interaction terms allow for heterogeneous slope coefficients. They contain only part of the control variables that enter directly.

III. Specification issues

The problems of estimating equation 10 are obvious: From equation 6 we know that the optimal choice of education depends on b_i and r_i. This leads to two potential endogeneity problems:

1. First, the individual effect a_i that represents unobserved ability might be correlated with b_i and r_i and thus observed education. This is the well-known *ability bias* (Griliches, 1977).

2. Second, $(b_i - \bar{b})$ might be correlated with observed education (again through b_i and r_i). This leads to a *self-selection bias* because people with higher returns to education might tend to acquire more education.

Both effects lead to an upward-biased estimate of returns to education (Card 1998).

Another important problem is the existence of reporting errors, both in wages and education. While the measurement error in wages only affects the estimation of the intercept, the error in reported education is more serious and might lead to an attenuation bias of the estimates of returns to education. Such a downward-bias would work in the opposite direction of the endogeneity bias described above and might offset those effects at least partially. Here we are concentrating on the possible solutions to the endogeneity problem. The problem of measurement error is not addressed here.[5]

In this paper we focus on family background as one possibility of mitigating the endogeneity problem. Factors subsumed under the notion „family background“ are for example parents' education, wealth, and family size as well as the existence of a complete or incomplete family.

One approach to reduce the bias in estimating returns to education is to add further control variables that might capture part of the unobserved components in the error term in equation 9. These controls should also enter in form of interaction terms with education to allow for

[5] The bias from measurement error might be substantial (see e.g. Ashenfelter and Krueger, 1994). For this reason we clearly have to find a solution to this problem in the future.

heterogeneous slope coefficients. The higher the correlation between the added variables and the unobserved components, the lower will be the remaining endogeneity bias. Obvious candidates for such control variables are family background characteristics (F_i) that most likely capture part of the unobserved ability, e.g. through genetics. An interaction term between education and family background can capture the effect of family background on returns to education. Then, the estimated equation looks as follows:

$$\log(y_i) = \alpha + \beta_1 S_i + \beta_2 X_i S_i + \beta_3 S_i^{\,2} + \gamma_1 E_i + \gamma_2 E_i^{\,2} + \delta X_i + \mu_1 F_i + \mu_2 F_i S_i + \varepsilon_i \qquad (11)$$

As Card (1998) has shown, adding family background as a control can potentially reduce the bias relatively to OLS, but it will not make it disappear completely, unless all unobserved ability components are captured by family background.

Another solution to the problem would be the estimation of equation 9 by an instrumental variable approach. An appropriate instrument generally has to satisfy two conditions:

1. It has to be strongly correlated with the „contaminated“ variable, i.e. education in our case,
2. it must not be directly correlated with the dependent variable (i.e. wages), which means that it has to be orthogonal to the unobserved ability components of the error term in the wage equation (strict exogeneity).

Instruments based on institutional factors, like minimum school leaving age, are particularly useful because the exogeneity assumption is justifiable for that kind of instruments. It is a well-established stylized fact that there is a strong positive correlation between the amount of education obtained by parents and their children. This correlation has been used to justify the use of parents’ education as an instrument. Even if family background has no independent effect on children’s earnings, it is not clear, however, whether family background is a good instrument for children’s education. In most cases, the estimate of the returns to education is still very likely to be upward-biased (Card, 1998). The bias of the IV-estimate is presumably bigger than the simple OLS estimate, which in turn is very likely to be bigger than the OLS estimate when family background is added as a control variable. The existence of measurement error in education reinforces the ordering of the estimates mentioned above (Card, 1998). Moreover, the conditions for IV estimation are much stronger with heterogeneous returns to education than in the usual applications of IV estimation (Card, 1998). Because of the very restricted reliability of IV estimates using family background as instruments we do not perform an instrumental variable estimation here.

The use of siblings data is a third way to overcome the endogeneity problem. It is based on the idea, that at least part of the unobserved heterogeneity is common to members of the same family. The difference in unobserved ability as well as its importance in the determination of education should be significantly lower *within* than *between* families. This „family fixed effect“ can be „differenced out“. This reduces the endogeneity bias relatively to the usual cross-sectional estimator. It has to be noticed, however, that differencing the data may exacerbate the problem of measurement error, since a large part of the true signal is „differenced out“.

In differenced form, the estimated equation of the fixed-effects regression looks as follows:

$$\begin{aligned}\log(y_i) - \log(y_j) = \beta_1(S_i - S_j) + \beta_2(X_iS_i - X_jS_j) + \beta_3(S_i^{\,2} - S_j^{\,2}) + \gamma_1(E_i - E_j) \\ + \gamma_2(E_i^{\,2} - E_j^{\,2}) + \delta(X_i - X_j) + \mu_2 F_h(S_i - S_j) + \upsilon \end{aligned}, \qquad (12)$$

where the indices i and j denote two siblings and h denotes the household that the siblings belong to. υ is the new error term. Actually we did not estimate this specification in differenced form. Instead, we include a “family fixed effect“ to allow for more than two siblings in one family (which increases efficiency). We also tested whether a random-effects specification might be more adequate.

IV. Description of Data and Variables

In this section, we describe the features of the data used in our analysis. The data are drawn from the *German Socio-Economic Panel* (GSOEP), a longitudinal study of private households in Germany. We use two different samples: first, a full sample, and second, a sub-sample of siblings.

1. The German Socio-Economic Panel (GSOEP)

The structure of the GSOEP is very similar to the American *Panel Study of Income Dynamics* (PSID) (see Wagner et al. 1993 for a more detailed description). The GSOEP was started in 1984 in West Germany with 5,921 households, 12,290 interviewed persons, and more than 2,000 children under age 17 (GSOEP West). In 1996, 4,445 households with 8,606 interviewed persons were left in the sample. In 1990 another 2,179 households containing 4,453 persons from East Germany were added to the panel (GSOEP East). Moreover, two samples of German immigrants were added in 1994 and 1995. Our analysis is based on the

West German sample, since the education system and the incentives to participate in schooling were very different in East Germany.

The interviewing procedure of the GSOEP is as follows: All household members who are at least 17 years old participate in the interviews. Children below age 17 also belong to the sample and are interviewed as soon as they reach the age of 17. The GSOEP keeps track of all persons who were a member of a sample household at some point in time. When a person of an “original“ household leaves a household and forms or enters a new household, this new household and all of its members are introduced into the sample. The GSOEP has been particularly successful in this respect. We exploit this feature of the GSOEP to construct a sibling sample in addition to the full sample consisting of all persons in the 13th wave of the GSOEP (1996).

The identification number of the biological mother was used to identify siblings. Thus, all persons in the sample who have the same mother are treated as siblings. Over time many of the children of a household reach age 17 and eventually enter the labor force. Now that the GSOEP has existed for 13 years an analysis based on siblings has become feasible with sample sizes that are reasonably large. There remain some problems with the sibling sample that are going to be discussed in detail below. We use all siblings contained in the data set and not just sibling pairs with exactly two siblings. Thus, the sibling sample can be viewed as an unbalanced panel.

We exclude persons who did not report a positive wage.[6] Since apprentices earn a reduced salary in Germany while they receive formal training and training on the job, we also have to exclude those from the working sample. Moreover, we exclude persons aged 57 and over since early retirement often starts at age 57. By excluding persons in the retirement age we avoid a bias that may arise if retirement decisions depend on wages.[7]

[6] This may introduce a selection bias. We do not deal with this problem here.

[7] This age restriction is not binding for the sibling sample.

2. Variable Description

The dependent variable in all regressions is the log of wage per hour (LOGWAGE) constructed from monthly gross income and the time actually worked per month.[8] Education (DEGREE) was constructed from data on the individuals' education degrees. Each degree enters with the number of years necessary for obtaining the degree and not with the time actually spent in an educational institution. We departed from that procedure in the case of vocational training (e.g. apprenticeships) where about half of the time is spent on the job and not in school. Thus, an apprenticeship, for example, was counted as one and a half years in the variable DEGREE instead of three years.

The data does not allow us to control for school quality. In Germany differences in school quality arise primarily because of different school types, while the quality differences within one school type are relatively small compared, for example, with the United States. There are three types of secondary schools: Secondary general school („Hauptschule“), intermediate school („Realschule“), and grammar school („Gymnasium“). There also is a distinction on the academic level between universities and vocational schools („Fachhochschulen“) that are more practically orientated than universities. One feature of the system is that the duration of study is also an indicator of school quality (e.g. secondary general school takes only nine years compared to 13 years in grammar school). Thus, the duration of education captures more than just the quantity aspect.

Table 1 shows the controls used in the basic specification. Age and the birth cohort are not separately identified in a cross-section. For this reason we use only one of them, namely the birth cohort (COHORT). The sufficiently exogenous variables SINGLEPARENT and SMALLKID*FEMALE are to capture the major reasons for people to work only part-time. CIVILSERVANT mainly captures the different tax treatment of civil servants.

Family background variables are briefly described in table 2. We measure family background in two different ways: We constructed variables similar to the DEGREE variable for the parents' years of education, and we also used dummies for parents' college degree. These variables appear to capture the influence of family background reasonably well. Including

---

[8] It is important to note that returns to education measured from hourly wage data are smaller than those from wages with a larger time frame, since hourly wages do not capture the effect of education on the time worked.

other variables like the family size or the age difference between parents and children did not change the results significantly and did not provide any further interesting insights.

Table 1: Controls used in the regressions.

EXPERIENCE	General working experience in years, constructed from biography data
TENURE	Job tenure = number of years with the current employer
COHORT	Birth cohort, deviation from sample mean
FEMALE	Dummy variable, 1 if female
SINGLEPARENT	Dummy variable, 1 if somebody is raising children without a partner
SMALLKID*FEMALE	Dummy variable, 1 if a woman has at least one child under age 16
CIVILSERVANT	Dummy variable, 1 if civil servant
Other variables:	Interaction terms of FEMALE and COHORT with DEGREE, squares of EXPERIENCE and COHORT

Table 2: Family background variables used in part of the regressions.

MOTHERDEGREE, FATHERDEGREE	Mother's/father's years of education, constructed in the same way as DEGREE, deviations from sample mean
MOTHERCOLLEGE, FATHERCOLLEGE	Dummy variable, 1 if mother/father has a college degree
PARENTSCOLLEGE	Dummy variable, 1 if father <u>or</u> mother has a college degree *)
Other variables:	Interaction terms of these variables with DEGREE

Notes: *) There are only very few cases in the sample where the mother has a college degree, but not the father.

3. Sample Characteristics

The working samples consist of 3,249 (full sample) and 495 persons (sibling sample). The size of the sibling sample is reduced strongly compared to the full sample, since a fixed-effects estimation requires at least pairs of observations with complete data. In the following we describe the main characteristics of, and differences between, these two samples. The descriptive statistics of the two samples including standard errors, minima, and maxima are shown in table 3.

One major difference between the two samples is the age of the sample persons. By construction, the sibling sample is much younger than the full sample, since the sibling sample is made up of children of panel households. Average ages are 29 years and 37 years, respectively. We observe only very few siblings above age 37. As a consequence of the difference in age distributions, the sample distributions of other variables differ as well: While the average durations of education are quite similar in the two samples (11.73 versus 11.41 years), mean experience (21.45 versus 14.33 years) and mean tenure (9.48 versus 5.74 years) differ markedly.

Table 3: Summary Statistics of the working samples.

	Full Sample (#obs = 3249)				**Sibling Sample** (#obs = 495)			
Variable	**Mean**	**Std. Dev.**	**Min**	**Max**	**Mean**	**Std. Dev.**	**Min**	**Max**
wage	25.51	16.37	3	451	22.92	13.65	3	258
degree	11.73	2.69	8	20	11.41	2.50	8	20
experience	21.45	12.85	0	55	14.33	8.24	0	41
tenure	9.48	8.54	0	41	5.74	4.59	0	24
cohort	58.94	9.76	40	78	67.17	4.81	49	77
female	0.42	0.49	0	1	0.32	0.47	0	1
motherdegree	8.68	2.41	0	18	8.75	1.67	0	18
fatherdegree	9.64	2.81	0	18	9.61	2.18	0	18
mothercollege	0.015	0.12	0	1	0.01	0.10	0	1
fathercollege	0.07	0.26	0	1	0.06	0.24	0	1
parentscollege	0.075	0.26	0	1	0.06	0.24	0	1

The distributions of gross monthly income of both samples have a first peak at the social security minimum threshold and a second peak at about 3,600 DM.[9] The means of gross monthly incomes are 4,234 DM in the full sample and 3,886 DM in the sibling sample. The level of wages is roughly 9 percent lower in the sibling sample than in the full sample. This can be attributed to the short labor market history of the persons in the sibling sample.

Family background variables are very similar in the two samples. In the full sample 1.5 percent of the mothers and 7 percent of the fathers have a college degree. Mothers' and fathers'

---

[9] Both samples exclude apprentices. The income of the sibling sample <u>including</u> apprentices shows a third peak at an income of 800 to 1,200 DM. This peak is due to the large number of persons in the sibling sample who are still apprentices and thus receive a reduced salary during their training (12 percent of that sample). If one excludes apprentices from both samples, the income distributions converge.

average duration of education are 8.68 and 9.64 years, respectively. In the sibling sample 1 percent of the mothers and 6 percent of the fathers have a college degree. Mothers' and fathers' average years of education are 8.75 and 9.61 years.

V. Estimation Results

First, we discuss the results from the regressions using the full sample. These results show the relevance of family background and the interesting interplay between level and slope effects. Heterogeneity seems to be an important feature of returns to education. Second, we present the results from the sibling sample, using the same specifications as in the full sample. These sibling regressions are to be understood mainly as benchmarks for the results from the fixed-effects estimation that provide further interesting insights. The marginal returns to education are calculated from equation 7, using point estimates of the coefficients and the sample means of all involved variables. Note that the reported returns are the ones for men. Since we normalized all variables that are interacted with education (except for the dummy variables) around their sample means, the returns can be read directly from the coefficients of education and its square. The regression output is presented for selected regressions.

1. Results for the Full Sample

Table 4 gives an overview of the estimated returns from the main regressions in the full sample. In table 5 we present the full regression results for our preferred specification (including the parents' college dummy and the corresponding interaction term).[10]

We first estimated a simple human capital earnings function similar to the standard Mincer equation (see equation 1). Most coefficients are highly significant and all have the expected sign. The estimated returns to education of 7.1 percent lie in the range of results reported in the German literature. The returns for women are about one percentage point higher than the ones for men.[11]

---

[10] The results on the other specifications are available upon request.

[11] This result also holds when the regressions are run separately for men and women.

Table 4: Marginal returns to education for different specifications in the full sample

	Marginal Returns to Education
Mincer equation	7.1% (0.38%)
Including parents' years of education	6.8% (0.64%)
Including parents' college dummy	7.0% (0.48%)
Including parents' years of education and the corresponding interaction terms	6.6% (0.52%)
Including the parents' college dummy and the corresponding interaction term	4.1% if PARENTSCOLLEGE = 1 (1.04%) 6.6% if PARENTSCOLLEGE = 0 (0.48%)

Notes: Standard errors in parantheses.

Adding parents' years of education as an additional regressor lowers the estimated returns to education slightly to 6.8 percent. Father's years of education appear to have a significant positive effect on the *level* of wages while mother's education is not significant. This is a result often reported in the literature using US data. The college dummy (PARENTSCOLLEGE) is not significant when added to the regression instead of parents' years of education. It becomes highly significant, however, when it enters not only linearly, but also as an interaction term with education (see table 5), and the null hypothesis that family background does not matter is rejected at the one percent level. While the parents' college degree has a positive impact on the *level* of wages, it has a negative effect on the *returns*. This result appears to be very robust and is consistent with the literature using US data (Ashenfelter and Rouse, 1998). Adding interaction terms lowers returns quite a bit: Returns are 6.6 percent if none of the parents has a college degree, but only 4.1 percent if at least one of the parents has a college degree. In the context of the above outlined theoretical model the interplay between returns and level might be interpreted as follows: Returns to education increase in both the marginal benefit parameter b_i and the marginal cost parameter r_i (see equation 7). Optimal schooling depends positively on the marginal benefit parameter b_i, but negatively on the marginal cost parameter r_i (see equation 6). The fact that we observe a strong positive correlation between parents' and children's education (see table 6), while returns are lower for children from families with higher education, suggests that lower marginal costs are the main reason for children from well-educated families attaining higher schooling in spite of lower returns. This is a very interesting and surprising result. In the German education system there are no tuition fees and rather generous support programs for children whose parents cannot afford to give them financial backing during their studies. Thus, one might think that credit market considerations play only a minor role in Germany. This means that taste factors and psychic costs - that surely both depend on family background – might play a very important role.[12]

Table 5: Regression results from the full sample including PARENTSCOLLEGE and its interaction term with education.

Depend. variable logwage	Coefficient	Std. Error	t-value	P>|t|
degree	0.1055	0.0238	4.42	0.00
$degree^2$	-0.0017	0.0009	-1.87	0.06
experience	0.0244	0.0031	7.80	0.00
$experience^2$	-0.0003	0.0001	-4.37	0.00
tenure	0.0046	0.0011	4.19	0.00
cohort	0.0227	0.0035	6.46	0.00
$cohort^2$	-0.0003	0.0001	-2.80	0.01
female	-0.2608	0.0701	-3.72	0.00
singleparent	-0.0629	0.0301	-2.09	0.04
smallkids*fem	-0.0515	0.0263	-1.96	0.05
civilservant	-0.0990	0.0227	-4.36	0.00
parentscollege	0.4042	0.1367	2.96	0.00
cohort*degree	-0.0018	0.0003	-5.71	0.00
female*degree	0.0098	0.0060	1.64	0.10
parentscollege*degree	-0.0253	0.0093	-2.71	0.01
constant	1.8383	0.1582	11.62	0.00

Notes: Heteroskedasticity consistent standard errors. Number of observations = 3249. R^2 = 0.3290.

Table 6: Correlation between children's and parents' education in the full sample.

	degree	fatherdegree	motherdegree
degree	1.00		
fatherdegree	0.43	1.00	
motherdegree	0.40	0.69	1.00

A similar interplay between linear terms and interaction terms can be observed for other variables like the sex dummy and the birth cohort: Women and earlier cohorts have a lower

[12] A very similar result has been found for the United States by Ashenfelter and Rouse (1998). For the U.S., this is far less surprising, however, since credit market considerations play a more important role.

level of wages, but a higher return. A Wald test confirms that heterogeneity of returns is present: The null hypothesis of homogeneous returns is easily rejected. These results suggest that a specification allowing for heterogeneous returns to education is indispensable for at least two reasons: first, it gives us interesting insights into the “mechanism” of the relationship between wage, education, and third factors; second, the omission of interaction terms might lead to inconsistent estimates. When parents’ education is measured in years instead of dummies, the effect disappears. Level and slope terms still have different signs, but they are insignificant.

2. Results for the Sibling sample

Now we present the results for the sibling sample that comprises 495 siblings from 224 different families (183 pairs, 34 triples, and 7 quadruples). We first present the results for the same specifications as in the full sample estimated by ordinary least squares. We do so in order to check how the results depend on the specific sample. Then, we present the fixed-effects estimates for the sibling sample.

Table 7 presents the regression results for the wage equation corresponding to table 5. In the sibling sample the estimates of the returns to education are lower than in the full sample. The coefficient of DEGREE is only 0.0319. The estimates of the marginal returns to education are much lower than in the full sample. Since the variables that are interacted with DEGREE are either centered around the sample mean or are dummies, the coefficient of DEGREE is readily interpreted as the marginal return to education for a male person with average birth year and parents without a college degree. The coefficient on PARENTSCOLLEGE*DEGREE is negative but insignificant. Thus, the point estimate of the returns on education for persons with college-educated parents are lower than for others, but not significantly different. This contrasts to the findings in the full sample, where children of college-educated parents seemed to have a significantly lower return to education. Similarly, the coefficient on PARENTSCOLLEGE has the same sign as in the full sample - namely positive -, but is also insignificant. The p-value for the test on joint significance of family background is 7 percent and thus the null cannot be rejected.[13] However, the hypothesis of homogeneous returns to education is rejected with a p-value of 1.6 percent.

---

[13] In specifications using other indicators for education of parents, the test yields similar results.

As for the full sample, we also tried other specifications. The returns to education are lower than in the full sample for all specifications of the wage function (see table 9). The result of heterogeneous returns is very robust as indicated by a Wald-Test on the joint significance of interaction terms.

Table 7: Wage regression for the sibling sample - OLS estimates.

Depend. variable logwage	Coef.	Std. Errors	t	P>|t|
degree	0.0319	0.0102	3.12	0.00
experience	0.0333	0.0081	4.09	0.00
experience²	-0.0007	0.0002	-2.62	0.01
tenure	0.0139	0.0037	3.76	0.00
cohort	0.0465	0.0219	2.13	0.03
female	-0.3127	0.1487	-2.10	0.04
singleparent	-0.1581	0.0681	-2.32	0.02
smallkids*female	-0.0544	0.0653	-0.83	0.41
civilservant	-0.0857	0.1027	-0.84	0.40
parentscollege	0.6774	0.5108	1.33	0.19
cohort*degree	-0.0039	0.0018	-2.13	0.03
female*degree	0.0176	0.0129	1.37	0.17
parentscollege*degree	-0.0296	0.0296	-1.00	0.32
constant	2.3363	0.1537	15.20	0.00

Notes: Heteroskedasticity consistent standard errors.
Number of observations = 495, number of families = 224, R-squared = 0.2376.
Test on heterogeneous returns (H_0: interactions are jointly zero) p-value = 0.016.
Test on family background (H_0: all family variables are jointly zero) p-value = 0.07.

The results between the full and the sibling sample can differ for several reasons. First, the sibling sample is much younger and thus experience and tenure differ, too. This may influence the estimates of returns to education if returns to experience depend on schooling - for instance it is very likely that the wage-experience profile is steeper for college graduates (see for instance Fitzenberger et al. 1997 for age-earnings profiles for German males). Also, the particular sample selection biases towards (i) persons with low education and (ii) towards persons who have completed their degree relatively quickly.

Second, by construction, the sibling sample has a higher degree of „within"-correlation, since it consists of siblings. Compared to the OLS-estimates in the full sample the within-variation in the sibling sample has more weight. This shifts the simple OLS-estimates towards the within-estimates.

In the next step we estimated the wage equation using a fixed-effects estimator for an unbalanced panel.[14] Simply introducing **family fixed effects** „explains“ 41 percent of the total variance in log-wages. Family background variables and other components of the individual-specific variation that are linearly related to family characteristics are absorbed by the family fixed effects. (As a matter of fact we cannot estimate the impact of those components in a fixed-effects approach). Family fixed effects also absorb heterogeneity in returns which is family specific. Only heterogeneity within the family is left. The results for the specification that corresponds to tables 5 and 7 are reported in table 8.

Table 8: Fixed effects wage regression for sibling sample.

Depend. variable logwage	Coefficients	Std. Errors	t-values	P>|t|
degree	0.0168	0.0156	1.08	0.28
experience	0.0344	0.0100	3.44	0.00
experience²	-0.0007	0.0003	-2.69	0.01
tenure	0.0141	0.0058	2.44	0.02
cohort	0.0732	0.0325	2.25	0.03
female	-0.3814	0.2157	-1.77	0.08
singleparent	-0.1891	0.1102	-1.72	0.09
smallkid*female	-0.1226	0.0940	-1.31	0.19
civilservant	0.0131	0.1115	0.12	0.91
cohort*degree	-0.0057	0.0027	-2.13	0.03
female*degree	0.0213	0.0184	1.16	0.25
parentscollege*degree	-0.0079	0.0165	-0.48	0.63
constant	2.5482	0.2325	10.96	0.00

Number of persons = 495, number of families = 224.
Standard deviation of family effects = 0.281, standard deviation of idiosyncratic effects = 0.335, total standard deviation of unobserved effects = 0.438, R-square within = 0.181.

[14] A Hausman test on Random Effects is rejected. We do not display the Random Effects estimates here, since they are of little interest given that they lie between OLS and Fixed Effects estimates.

The fixed effects estimate for the linear education term is only 0.0168 and is insignificant. Recall that the corresponding OLS estimate was 0.0319 in the siblings regression. In the full sample the OLS estimate was 0.1055 for the linear and 0.0017 for the squared term of education.. The total marginal return to schooling for our basic case (male, mean cohort, parents without college) is 1.68 percent. For younger cohorts, the return is even lower. For children of college educated parents the point estimate of marginal returns to education is almost zero. Returns are still heterogeneous even if we control for family fixed effects although the effect is less pronounced. The test on homogeneous returns rejects the null hypothesis at a significance level of 7.4 percent.

In table 9 we display the point estimates for the marginal returns for different samples, estimators, and specifications. The last two rows correspond to the estimates reported in tables 5, 7, and 8. The basic findings are that the returns to education may be considerably overstated by OLS estimates of standard Mincer-type wage equations. Introducing family background controls and interaction terms to control for heterogeneous returns reduces the estimates of returns to education. Using the siblings subsample and fixed effects reduces the returns further. For instance, the fixed effects estimate of the Mincer equation gives a marginal return of only 2.4%, which is barely significantly different from zero. However, we have to keep in mind the potential bias arising from measurement error.

Table 9: Comparison of marginal returns to education for different specifications.

	Marginal Returns to Education		
	Full Sample OLS	Sibling sample OLS	Sibling sample Fixed Effects
Mincer equation	7.1% (0.38%)	4.4% (0.83%)	2.4% (1.37%)
Including parents' years of education	6.8% (0.64%)	3.5% (1.09%)	–
Including parents' college dummy	7.0% (0.48%)	3.9% (0.86%)	–
Including parents' years of education and the corresponding interaction terms	6.6% (0.52%)	2.3% (1.08%)	3.1% (1.56%)
Including the parents' college dummy and the corresponding interaction term:			
- PARENTSCOLLEGE = 1	4.1% *) (1.04%)	0.2% (2.85%)	0.9% (1.95%)
- PARENTSCOLLEGE = 0	6.6% (0.48%)	3.2% (1.02%)	1.7% (1.56%)

Notes: Standard Errors in Parantheses. *) If we evaluate the marginal returns for the full sample at the means of the sibling sample, the marginal returns are lower (2.64% and 5.17), but still well above the results for the siblings regressions.

VI. Conclusion and Extensions

In this paper we examined the role of family background in the determination of wages and its impact on returns to education. The results from the full sample suggest that family background in fact does play an important role in the determination of wages. Moreover, we found that returns to education seem to be heterogeneous and influenced by family background variables. Children of less educated parents have lower wages, but higher returns to education with the correlation between children's and parents' schooling being positive. According to the theoretical model this result might be due to higher marginal costs of children in families with lower educational levels. Since credit market considerations play a minor role in Germany, taste differences and psychic costs may play a major role.

The estimates from the full sample suffer, however, from the problem of endogeneity of education. Thus, we have used a sibling analysis in order to reduce this bias. The results from our analysis confirm the presumption that the OLS estimates might overstate the true returns to education. On the other hand, measurement error might lead to a downward bias in the estimates. Further research hopefully helps us to approximate the size of these effects. We also found that household variation constitutes an important part of the variation in the logarithm of wages, which confirms the important role of family characteristics.

There are some caveats and possible extensions to our analysis. One central point is the neglect of the problem of measurement error in education and in family background variables. The fixed-effect estimation might exacerbate this problem in reducing the signal-to-noise ratio by „differencing".

Second, so far we have not exploited the panel structure of the data. Even though the panel does not help us to identify returns to education by differencing over time (since education itself would be differenced out), we could use the longitudinal aspect of the data set in order to increase efficiency or to generate additional identifying information, e.g. to solve the problem of measurement error.

Third, other specifications could be tested that allow for a more flexible form. For example, higher-order polynomials in education, experience, and tenure could be used as well as interaction terms between education and age or experience. More importantly, we should try a correlated random-effects analysis in order to examine the assumptions underlying our fixed-effects specification.

Another caveat concerns the adequacy of the siblings analysis for reducing the endogeneity bias. We might worry that the siblings are not really comparable, because for example the first child is treated differently than the second child. In fact, the first siblings' wages are higher than those of the second siblings. The same is true for education. Probably, most of this is due to the difference in age between the first and the second sibling. It might, however, also be a potential source of bias. As a first check, we included a dummy variable for being the first sibling into our regressions. The coefficient of this dummy was not significantly different from zero, which might indicate that there is no serious bias from this kind of problem. We plan to examine this problem in more detail. A similar effect could be present in the comparison of men and women. It might be good to use only sibling pairs of the same gender to avoid such a problem. Of course, this would hugely reduce the sample size.

As a last point, we want to mention that our analysis allows us to control for individual characteristics that are constant within families, but not for idiosyncratic ability factors. One might want to include for example an IQ measure in order to capture at least part of these factors. That kind of measure is not contained in the GSOEP, however. Moreover, it is questionable what these IQ measures really indicate.

References:

Abraham, K. G., and S. N. Houseman (1995): "Earnings Inequality in Germany," in: *Differences and Changes in Wage Structures*, R. B. Freeman, R. B. and L. F. Katz (eds.), Chicago and London: University of Chicago Press.

Altonji, Joseph G., and Thomas A. Dunn (1996): „The Effects of Family Characteristics on the Return to Education,“ *Review of Economics and Statistics*, 78: 692-704.

Ashenfelter, Orley, und Alan Krueger (1994): "Estimating the Returns to Schooling Using a New Sample of Twins," *American Economic Review*, 84, 1157-1173.

Ashenfelter, Orley, and Cecilia Rouse (1998): "Income, Schooling, and Ability: Evidence From a New Sample of Identical Twins," *Quarterly Journal of Economics*, 113, 253-284.

Becker, Gary S. (1967):“Human Capital and the Personal Distribution of Income,“ University of Michigan Press, Ann Arbor: Michigan.

Card, David (1995): „Earnings, Education, and Ability Revisited,“ Research in Labor Economics, volume 14, Solomon Polachek, editor, JAI Press: Greenwich Connecticut: 23-48.

Card, David (1998): „The Causal Effect of Education on Earnings,“ forthcoming in the *Handbook of Labor Economics*, Orley Ashenfelter and David Card (eds.).

Fitzenberger, B., R. Hujer, Th. MaCurdy, and R. Schnabel (1995): "The Dynamic Structure of Wages in Germany - A Cohort Analysis," Discussion paper 533-95, University of Mannheim, Department of Economics.

Griliches, Zvi (1977): "Estimating the Returns to Schooling: Some Econometric Problems," *Econometrica*, 45, 1-22.

Licht, Georg, and Viktor Steiner (1991): "Stichprobenselektion, unbeobachtete Heterogenität und Humankapitaleffekte bei der Schätzung von Einkommmensfunktionen mit Paneldaten," in *Lebenslagen im Wandel: Zur Einkommensdynamik in Deutschland seit 1984*, Ulrich Rendtel and Gert Wagner (eds.), Frankfurt am Main: Campus.

Mincer, Jacob (1974): *Education, Experience and Earnings*, Columbia University Press: New York.

Nickell, S., and B. Bell (1996): "Changes in the Distribution of Wages and Unemployment in OECD Countries," *AER, Papers and Proceedings*, 86(2), pp 302-308.

Steiner, Viktor, and Kerstin Wagner (1996): "Has Earnings Inequality Changed in the 1980’s?" Center for European Economic Research, Discussion Paper 96-32, Mannheim.

Wagner, Gert, Richard Burkhauser, and Friederike Behringer (1993): "The English Language Public Use File of the German Socio-Economic Panel Study," *The Journal of Human Resources*, 28(2), 429-433.