Discussion Paper No. 02-67

Family and Gender Still Matter: The Heterogeneity of Returns to Education in Germany

Isabel Schnabel and Reinhold Schnabel

ZEW

Zentrum für Europäische Wirtschaftsforschung GmbH

Centre for European Economic Research

Discussion Paper No. 02-67

Family and Gender Still Matter: The Heterogeneity of Returns to Education in Germany

Isabel Schnabel and Reinhold Schnabel

Nontechnical Summary

In the 1960s, substantial reforms took place in the German educational system. The reforms aimed at equalizing the educational opportunities across social classes and gender. Important reforms were the abolition of tuition fees in schools and universities, the introduction of a means-tested income support (BAFöG), the building of new secondary schools and colleges. These measures reduced the individual costs of higher education considerably and led to an increase in educational participation during the 1970s and 1980s. The rise in the educational participation of women was particularly strong. However, the relative participation of children from poorer social backgrounds in higher education remained low, since the participation of the middle class also expanded. Thus, the link between family background and participation in higher education is still close.

This raises the question as to how family background determines wages and returns to education in Germany. We try to answer this question by examining the impact of family background on wages and on returns to education. In estimating the earnings function, we (i) introduce family background variables as control variables, (ii) allow for heterogeneous returns to education, and (iii) construct a siblings sample from the German Socio-Economic Panel (GSOEP). Our approach allows to control for unobserved family-specific heterogeneity and to eliminate the bias due to family effects. We use a fixed-effects estimator and, as an alternative, a correlated random-effects estimator.

Our main result is that family background still matters despite the attempts – or political claims – to equalize educational opportunities. Family characteristics constitute an important part of the variation in wages and in marginal returns to education, which confirms the important role of family background. Persons with well-educated parents tend to have lower returns to education and earn higher wages than persons with less-educated parents. Based on our theoretical model, we argue that this must be due to lower marginal costs of education in well-educated families. In addition, we find that gender matters for returns to education. While women, on average, earn lower wages than men, they have higher marginal returns to education. This may be explained by a self-selection of the more productive women into paid work – as opposed to household production. As a by-product of our analysis, the same interplay of wage levels and marginal returns was found for the effect of cohort membership on wages and marginal returns. The reduction in private marginal costs of education in the 1970s and 1980s has increased participation in schooling and thereby has reduced private returns. At the same time, macroeconomic growth has lead to an increase in the wage level of younger cohorts.

The reforms of the education system benefited above all the middle class. If public policy intends to increase the participation of children from poorer families, it will be necessary to lower their private marginal costs. Financial grants or loans may be one way to achieve that goal. The trend of restricting financial assistance to students is counterproductive as it leads to increased social selection and to an under-utilization of human resources.

Family and Gender Still Matter:

The Heterogeneity of Returns to Education in Germany

Isabel Schnabel* **and Reinhold Schnabel****

*Department of Economics, University of Mannheim

**Department of Economics, University of Essen, and ZEW, Mannheim

Revised September 2002

Abstract
Using information on family background, we estimate returns to education, allowing for the heterogeneity of returns. In order to control for the unobserved heterogeneity shared by family members, we construct a siblings sample and employ family fixed-effects and family correlated random-effects models. Our main result is that family background still matters despite the substantial political efforts to equalize educational opportunities in Germany. Persons with less-educated parents earn lower wages, but have higher returns to education. This supports the view that persons from less-educated backgrounds still face higher marginal costs in the educational system. The same interplay between the wage level and marginal returns is found for the effect of gender and cohort.
(JEL: J31, C23)
Keywords: Returns to education, siblings analysis, heterogeneity of returns.

Support from *Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 504* at the University of Mannheim and from *Deutscher Akademischer Auslandsdienst* (DAAD) are 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 received helpful comments from Axel Börsch-Supan and Bernd Fitzenberger.

Isabel Schnabel
Department of Economics
University of Mannheim
68131 Mannheim, Germany
Telephone: +49-621-181 1835
Fax: +49-621-181 1837
isabel@econ.uni-mannheim.de

Reinhold Schnabel
Department of Economics
University of Essen
45117 Essen, Germany
Telephone: +49-201-183 4125
Fax: +49-201-183 3417
reinhold.schnabel@uni-essen.de

Family and Gender Still Matter:

The Heterogeneity of Returns to Education in Germany

I. Introduction

In the 1960s, substantial reforms took place in the German educational system. The reforms aimed at equalizing the educational opportunities across social classes and gender. Important reforms were the abolition of tuition fees in schools and universities, the introduction of a means-tested income support for children in higher education (BaföG), the building of new secondary schools and colleges, and the attempts to equalize school quality throughout the country. These measures reduced the individual costs of higher education considerably. As a consequence, the participation in higher education rose dramatically during the 1970s and 1980s, as completion of secondary school and college enrollment grew rapidly. The rise in the educational participation of women was particularly strong. However, the relative participation of children from poorer social backgrounds in higher education remained low, since the participation of the upper middle class also expanded. Thus, the link between family background and participation in higher education is still close.

This raises the question as to how family background determines wages and returns to education in Germany. We will try to answer this question by examining the role of family background in the human capital earnings function, i.e., its direct influence on wages and its impact on returns to education. While this area of research has been very active in the United States in recent years, empirical work in Germany has focused on questions of industry-specific wage structure, wage dispersion, inequality, and wage stickiness (Abraham and Houseman 1995, Fitzenberger et al. 2001, Nickell and Bell 1996). Licht and Steiner (1991) estimate human capital earnings functions and try to control for unobserved heterogeneity using panel data. They rely on individual changes in education over time to identify the returns to education. However, this approach is likely to pick up measurement error, and moreover, treats persons with uncompleted education histories as if the transitory education status were optimal. Ichino and Winter-Ebmer (1999) use parental information as instruments to identify the effect of schooling on wages. Lauer and Steiner (2000) conduct a descriptive analysis of the wage-education relation by estimating a standard earnings function disaggregated by cohort and age groups.

The estimation of marginal returns to education is plagued by problems of unobserved heterogeneity. Unobserved ability tends to bias the estimates of returns to education upwards; this is the "classical" ability bias (Griliches 1977). Moreover, the "self-selection" of persons with high ability into higher education introduces heterogeneity into the return to education. This biases the conventional estimates of returns to schooling upwards (Card 1999, 2001). In order to reduce the bias in the estimation of returns to education, we (i) introduce family background variables as additional control variables, (ii) allow for heterogeneous returns to education, and (iii) construct a siblings sample from the German Socio-Economic Panel (GSOEP). Such an analysis has not been done for Germany before. Our approach allows to control for unobserved family-specific heterogeneity and to eliminate the bias due to family effects. We use a fixed-effects estimator and, as an alternative, a correlated random-effects estimator.[1] Since the fixed-effects estimator is particularly sensitive to measurement errors in the education variable, we favor the correlated random-effects estimator.[2]

Our empirical results can be summarized as follows: In the OLS regression based on the full sample, family background plays an important role in determining wages. Moreover, returns to education are heterogeneous, with family background being an important source of this heterogeneity. Persons from less-educated families have lower wages, but also higher returns. Since the OLS estimates of returns to education are likely to be biased upwards, we undertake a siblings analysis. As expected, the family fixed-effects estimates from the siblings sample are smaller than the OLS estimates, with marginal returns close to zero. We find that family-specific effects account for the larger part of the variation in wages. Using the correlated random-effects estimator as an alternative to the fixed-effects estimator, we confirm the findings that returns to education are heterogeneous and vary by gender, cohort, and family background. Women display higher returns to education than men, but lower wage levels. Older cohorts also display higher returns and lower wages than younger cohorts. Most importantly, children from less-educated families have higher marginal returns to education than children of highly educated families. Based on our theoretical model, this implies that their marginal costs of education are higher, while we can rule out the hypothesis that their marginal benefit function is shifted to the left. Despite the substantial political efforts to

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

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

equalize educational opportunities in Germany, family background still is an important determinant of education and wages.

The remainder of the paper is organized as follows: We start with a brief description of the German educational system. In section III, we derive the economic model as the basis of our empirical analysis. In the next section, we discuss specification issues and solutions to the endogeneity problem. A description of the data set and the variables is found in section V. Section VI presents the result. First, we show the estimated earnings function based on the full sample, controlling for heterogeneity in wage levels and returns to education by using family background variables. Then we present the results for the siblings sample, using fixed-effects and correlated random-effects estimators. Section VII concludes.

II. Schooling and Education in Germany

Germany features a compulsory schooling of nine years starting at the age of six. During the first four years, all children have to attend a primary school in their neighborhood. Since the financing of primary schools is uniform across neighborhoods and more or less so across states as well, the main reasons for differences in quality of primary schooling are neighborhood and peer effects. Due to social and ethnic segregation, the effective quality may vary considerably across primary schools. At the end of the fourth year, a decision on the form of secondary education has to be made, since Germany has a *three pillar secondary-schooling system,* which starts with the fifth year. Depending on the performance in primary school, children are selected into either a “basic” secondary school (5 years, “Hauptschule”), an intermediate secondary school (6 years, “Realschule”), or a grammar school (9 years, “Gymnasium”).[3] Parents are free to choose secondary schools outside their school district, and even outside their state. Each of the three types of secondary schools are usually completed with graduation. The quality differences within one school type are relatively small compared, for example, to the United States. The quality differences across school types are huge.

After secondary school, the majority starts a (modestly) paid apprenticeship in a firm, which typically lasts between two and three years. This is a practical training in the workplace and in special employee-sponsored courses, augmented by a more theoretical training in state

---

[3] Some States introduced “comprehensive secondary schools” as a voluntary alternative to the three-pillar system. This experiment was not successful due to the massive self-selection into traditional grammar schools.

schools (usually one day per week). The apprentices spend about half of their time in training and finish their education with a formal degree. Following the apprenticeship, there is a variety of further qualification steps that may eventually lead to admission in an applied college ("Fachhochschule"). Admission to university generally requires graduation from grammar school. Applied colleges usually require three years of course work, while universities require between four and five years of course work and have a more theoretical focus.

After the reforms in the 1960s, the participation in secondary and tertiary education rose quickly. In 1970, 60 percent attended the five-year secondary school, while only 22 percent attended a grammar school. In 1990, only 40 percent chose the five-year secondary school, while one third went to a grammar school.[4]

Participation in college education had been steady until the end of the 1950s. For decades, the college graduate rate stood at about 8 percent for males and at barely 3 percent for females. Participation in college education picked up slightly during the 1960s, and rose considerably after the reforms. Corrected for cohort size, the number of students increased by 70 percent from 1970 to 1980, and increased by another 37 percent by 1990 (Wissenschaftsrat 2002).[5] In 1980, 15 percent of the males at the age of 30 to 34 had earned a college degree. Participation of women in college education rose even more strongly. By 1980, college attendance of females had more than tripled compared to the numbers in the 1950s, and 10 percent of females aged 25 to 29 held a college degree. In 1980, 41 percent of those who enrolled in college were females.

Educational attainment is still highly correlated with family background. Only 12 percent of the working-class children enroll in a college, while the enrollment rate is almost three quarters for children from civil-servant households (Wissenschaftsrat 2002). Also, 75 percent of children whose father has a college degree also attend a college. During the last decade, the link between family background and education has become even stronger. One reason for this reversal might be that income support programs for students have been scaled down since the early 1980s through a partial switch from grants to loans combined with stricter means tests.

---

[4] The numbers on educational attainment are own calculations based on Wissenschaftsrat 2000, IW 2002, and on the data of the 1% census surveys from 1976 to 1995 (see also Fitzenberger, Schnabel, Wunderlich 2001).

[5] In international perspective, college enrollment rates are still relatively low. Only 20.4 % of the relevant cohort in 1980 and 27.3 % in 1990 enrolled in tertiary education. Moreover, college drop-out rates are quite high at about 40%; as a consequence, only 16% of a cohort completed a college degree (Wissenschaftsrat 2002).

These changes have increased the cost of college education for children from low-income families (Schimpl-Neimanns 2000).

III. An Empirical Model of Returns to Education

The research on returns to education started with the work by Becker (1967) and Mincer (1974), who introduced the *human capital earnings function*, which still is the basis of practically all research on returns to education. In the most basic specification, returns to education are estimated as follows:

$$\log(y_i) = \beta_0 + \beta_1 S_i + \beta_2 E_i + \beta_3 E_i^{\,2} + \varepsilon_i \, , \tag{1}$$

where y_i is a measure of income, S_i is years of completed education, E_i is experience, and ε_i is a statistical residual. This basic specification has been extended by economists in several directions to capture additional aspects of returns to education. The framework used in this paper was introduced by Card (1995, 1999)[6]. 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 life-time utility, which depends on the average level of earnings over the lifecycle, denoted by y(S), and the disutility of education, h(S):

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

Linearity in log earnings implies that the optimal education choice does not depend on factors that raise earnings proportionally for *all* levels of education. $h(S_i)$ is an increasing and convex function of S_i. The first-order condition of this maximization problem sets the marginal benefits from education equal to its marginal costs:

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

Individual heterogeneity is introduced 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:

---

[6] This section follows the presentation in Card (1999).

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

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

$$h'(S_i) = r_i + k_2 S_i \qquad (5)$$

The optimal level of education derived from the first-order condition (3) is:

$$S_i{}^* = \frac{b_i - r_i}{k_1 + k_2} \qquad (6)$$

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

$$\beta_i = b_i - k_1 S_i{}^* = b_i\left(1 - \frac{k_1}{k_1 + k_2}\right) + r_i \frac{k_1}{k_1 + k_2}, \qquad (7)$$

which simply is a convex combination of the two intercept terms from equations 4 and 5. 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., $b_i = \bar{b}$ $\underline{and}$ $k_1 = 0$) or
2. marginal costs are constant and equal for all i (i.e., $r_i = \bar{r}$ $\underline{and}$ $k_2 = 0$).

By integrating equation 4, we can derive the following model:

$$\log(y_i) = a_i + b_i S_i - \frac{1}{2} k_1 S_i{}^2 \qquad (8)$$

The important differences to the Mincer model in equation (1) are the individual-specific intercept and slope terms, which turn the model into a *random-coefficient model*. The equation can be written in the form of deviations from means in the following way:

$$\log(y_i) = \bar{a} + \bar{b} S_i - \frac{1}{2} k_1 S_i{}^2 + (a_i - \bar{a}) + (b_i - \bar{b}) S_i \qquad (9)$$

IV. Econometric Model and Specification Issues

Many empirical studies on returns to education assume that the constant and the coefficient of education in the wage equation are constant across individuals, implying that the last two terms in equation (9) enter the error term. This potentially leads to two endogeneity problems:

1. The individual effect a_i, which represents unobserved ability, might be correlated with b_i and r_i and, thus, with observed education (cf. equation 6). This is the well-known *ability bias* (Griliches 1977).
2. The last term in equation (9) is correlated with observed education if returns to education are not equal across individuals. The resulting bias has been named *self-selection bias* because it arises from people with higher returns who wish to acquire more education.

Under weak assumptions, both effects lead to an upward-biased estimate of returns to education (Card 1999).

Another important problem may be the existence of measurement errors in wages and education. While the measurement error in wages primarily affects the estimation of the intercept, the measurement error in education is more serious, because it leads to an attenuation bias of the estimates of returns to education.[7] The bias works in the opposite direction of the endogeneity bias described above and might offset those effects at least partially. Measurement errors in education may be due to reporting errors or due to the fact that the commonly used education variables neglect the quality of education.

An econometric model that takes into account the heterogeneity of returns to education can be formulated as follows, where X_i denotes all control variables besides the linear education term:

$$\log(y_i) = \beta_0 + (\beta_1 + \beta_{2i}) \cdot S_i + X_i \beta_3 + \alpha_i + \varepsilon_i , \qquad (10)$$
$$\text{where} \quad E(\alpha_i \mid S_i) \neq 0, \quad E(\varepsilon_i \mid S_i, X_i) = 0 .$$

α_i denotes the part of a_i that is correlated with education (the so-called *individual effect*), while ε_i denotes an idiosyncratic error term. If α_i was observed, including it into the regression would solve the endogeneity problem, but this is usually not an option because “ability” is mostly unobserved. The heterogeneity of returns can be taken into account by modeling the returns themselves as a function of observables, leading to the following model:

$$\log(y_i) = \beta_0 + (\beta_1 + X_i \beta_2) \cdot S_i + X_i \beta_3 + \alpha_i + \eta_i S_i + \varepsilon_i , \qquad (11)$$
$$\text{where} \; E(\alpha_i \mid S_i) \neq 0, \quad E(\eta_i \mid S_i) \neq 0, \; E(\varepsilon_i \mid S_i, X_i) = 0.$$

The part of the heterogeneity in returns that cannot be explained by observables enters the error term.

[7] The bias from measurement error may be substantial (see, e.g., Ashenfelter and Krueger 1994).

In this paper, we focus on the use of family background information as one possibility of mitigating the endogeneity problem. We use parents' education as a proxy for family background. One way to reduce the bias in estimated returns to education is to add further control variables that capture part of the unobserved components contained in the error term. These control variables should also enter in the form of interaction terms with education to allow for heterogeneous slope coefficients. The higher the correlation between the added variables and the unobserved components, the lower the remaining endogeneity bias. Family background variables are obvious candidates for such control variables because they are most likely to capture part of the unobserved ability, e.g., through genetics. However, adding family background as control variables will not make the bias disappear completely, unless all unobserved ability components can be captured by family background.

Another solution to the endogeneity problem is the estimation of equation (11) by an instrumental variables approach. Instruments based on institutional factors, like minimum school leaving age, are particularly useful because the required exogeneity assumption is justifiable for that kind of instrument. The strong positive correlation between the years of education obtained by parents and their children has been used to justify the use of parents' education as an instrument. However, even if family background has no independent effect on children's earnings, it is not clear whether family background is a good instrument for children's education. As Card (1999) has shown, the estimated returns to education are still upward-biased in most cases. The bias of the IV-estimate is presumably larger than the simple OLS estimate, which in turn is very likely to be larger 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 1999). Moreover, the conditions for IV estimation are much stronger with heterogeneous returns to education than in the usual applications of IV estimation (Card 1999). Because of the restricted reliability of IV estimates using family background as instruments, we will 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. The model from equation (11) is written as follows:

$$\log(y_{sf}) = \beta_0 + (\beta_1 + X_{sf}\beta_2 + X_f\beta_3)\cdot S_{sf} + X_{sf}\beta_4 + X_f\beta_5 + \alpha_{sf} + \eta_{sf}S_{sf} + \varepsilon_{sf}\,, \qquad (12)$$

where s = 1,... S denotes the sibling and f = 1,... F the family. The X's not indexed by s are variables that are common to members of the same family. Let us split up the individual effect α_{sf} into two components, such that $\alpha_{sf} = \alpha_f + (\alpha_{sf} - \alpha_f)$, where α_f denotes the "*family effect*":

$$\begin{aligned}\log(y_{sf}) = \beta_0 + (\beta_1 + X_{sf}\beta_2 + X_f\beta_3)\cdot S_{sf} + X_{sf}\beta_4 + X_{sf}\beta_5 \\ + \alpha_f + (\alpha_{sf} - \alpha_f) + \eta_{sf}S_{sf} + \varepsilon_{sf}\end{aligned} \qquad (13)$$

Differencing the model across siblings (or, alternatively, including family dummies), leads to

$$\Delta\log(y_{sf}) = (\beta_1 + X_f\beta_3)\cdot\Delta S_{sf} + \beta_2'\Delta(X_{sf}'S_{sf}) + \Delta X_{sf}\beta_4 + \Delta\alpha_{sf} + \Delta(\eta_{sf}S_{sf}) + \Delta\varepsilon_{sf} \quad (14)$$

In the case where both α_{sf} and η_{sf} are the same within families (a so-called "*pure family-effects model*"), the endogeneity bias would disappear completely. In the more realistic case where α_{sf} and η_{sf} are not the same, estimated returns to education will still be biased. However, the bias will be reduced compared to the usual cross-sectional estimator if the educational choice *within* families depends less on ability than the choice of schooling in the population.[8]

One problem in using a fixed-effects procedure as in equation (14) is the large number of parameters to be estimated, leading to relatively large standard errors. Moreover, the problem of measurement error might be exacerbated as the signal-to-noise ratio is reduced by "differencing out" a large part of the true signal. A smaller education coefficient in a fixed-effects regression compared to OLS might not reflect a smaller endogeneity bias, but simply an attenuation from measurement error. Therefore, we also estimate a more general model that explicitly models the correlation between the unobserved individual effect and the regressors. Consider the wage equation (13) for sibling s = 1,... S in family f = 1,... F and write it in a more compact and simplified way as

$$\log(y_{sf}) = X_{sf}\beta + \alpha_f + \varepsilon_{sf}\,, \qquad (15)$$

where $E(\varepsilon_{sf} \mid X^S{}_f, \alpha_f) = 0\ \forall s, f$. (16)

[8] In our study, we do not estimate the model 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 test whether a random-effects specification might be more adequate than a fixed-effects specification.

X_{sf} denotes here a 1×K-vector of all regressors that vary across siblings (including education)[9], while X^S_f is a 1×(S·K)-vector containing the X's of *all* S siblings in family f. We assume that the (latent) family effect α_f fully captures the individual effect, i.e., $\alpha_{sf}=\alpha_{s'f}=\alpha_f$ for all siblings from the same family, and that the heterogeneity of returns can be fully modeled through observed variables. Note that assumption (16) requires the X's to be strictly exogenous, while it allows for a correlation between the family effect and the X variables.

In the context of panel data, Chamberlain (1982, 1984) proposes a generalization of the random-effects model that allows the conditional expectation of the individual effect to be an unknown function of the regressors from *all* waves of the panel. Translated into our context, this could be written in the following way:

$$E(\alpha_f \mid X^S{}_f) = X^S{}_f \lambda = X_{1f}\lambda_1 + X_{2f}\lambda_2 + ... + X_{Sf}\lambda_S \qquad (17)$$

Equation (17) states that the expectation of the family effect, conditional on the X's from all waves, is a *linear* function of these X's. In fact, this linearity assumption could easily be relaxed by using a (minimum mean-squared-error) linear projector instead of the conditional expectation (cf. Chamberlain 1982, 1984). Combining equations (15), (16), and (17), we can write our model as

$$E[\log(y_{sf}) \mid X^S{}_f] = X_{sf}\beta + X^S{}_f \lambda, \qquad (18)$$

$$\text{where} \quad \nu_{sf} = \log(y_{sf}) - E(\log(y_{sf}) \mid X^S{}_f) \qquad (19)$$

$$\text{and} \quad E(\nu_{sf} \mid X^S{}_f) = 0. \qquad (20)$$

The model in equation (18) contains (1+S)·K coefficients (plus the variance parameters). Chamberlain (1982) has shown that efficient estimation can be done by applying a ***two-step minimum distance*** procedure: In a first step, the model from equation (18) is estimated in its *reduced form* by a simple unconstrained OLS regression separately for each group of siblings (i.e., the group of "first" siblings, "second" siblings etc.[10]):

$$\log(y_{sf}) = X^S{}_f \pi_s + \nu_{sf}\text{, s = 1,... S} \qquad (21)$$

[9] We neglect for the moment those variables that do not vary across siblings for the ease of the exposition.

[10] Note that the "first" ("second", "third",...) sibling is the sibling first born in a family, not the oldest sibling in our sample. Thus, λ_s is restricted to be the same for all first-born siblings (second, third,...).

where π_S is a (S·K)×1-vector of reduced-form coefficients. In a very compact form, the complete system of equations can be written as

$$\underset{SF\times 1}{\log(y)} = [\underset{S\times S}{I_S} \otimes \underset{F\times SK}{X^S}] \cdot \underset{S^2K\times 1}{\pi} + \underset{SF\times 1}{\nu} \ . \tag{22}$$

This system of equations is a special case of a SUR model with identical explanatory variables, such that an OLS estimation equation by equation is identical to a GLS estimation of the whole system (Zellner 1962). The number of coefficients to be estimated in this first step is $S^2 \cdot K$, which far exceeds the number of structural coefficients $(1+S)\cdot K$. In order to account for the unbalanced panel, all rows of X referring to "non-existent" siblings (in the sample) must be filled up with zeros.

The covariance matrix is modelled flexibly in that we allow the variances of the errors and their covariances *within* families to differ across families, while we restrict the covariances *across* families to zero. In the spirit of White (1980), the asymptotic covariance matrix of the reduced form coefficients can be estimated using the estimated residuals from the OLS regression[11]:

$$\underset{S^2K\times S^2K}{\hat{\Omega}} = \left(I_S \otimes \left(\frac{1}{F} \sum_f X^S{}_f{}' X^S{}_f \right)^{-1} \right) \cdot \left(\frac{1}{F} \sum_f \left(\hat{u}_f \hat{u}_f{}' \otimes X^S{}_f{}' X^S{}_f \right) \right)$$
$$\cdot \left(I_S \otimes \left(\frac{1}{F} \sum_f X^S{}_f{}' X^S{}_f \right)^{-1} \right). \tag{23}$$

In a second step, the structural parameters are estimated applying a minimum-distance procedure, minimizing the objective function

$$\left(\hat{\pi} - f(\beta,\lambda)\right)' \cdot \hat{\Omega}^{-1} \cdot \left(\hat{\pi} - f(\beta,\lambda)\right), \tag{24}$$

where the optimal weighting matrix is the inverse of the covariance matrix from step 1 and where $f(\beta,\lambda)$ is the $(S^2K)\times 1$-vector of parameter restrictions

$$f_{ss'}(\beta,\lambda) = \left[1\{s = s'\} \cdot \beta + \lambda_{s'}, s, s' = 1,\ldots S\right]. \tag{25}$$

The covariance matrix of the structural coefficients $\hat{\theta}$ can be estimated consistently by

$$est.Var(\underset{K(T+1)\times 1}{\hat{\theta}}) = \left(\frac{\partial f(\hat{\theta})'}{\partial \theta} \cdot \hat{\Omega}^{-1} \cdot \frac{\partial f(\hat{\theta})}{\partial \theta'} \right)^{-1}. \tag{26}$$

---

[11] More precisely, Ω is the asymptotic variance of $\sqrt{F} \cdot (\hat{\pi} - \pi)$.

The minimum distance procedure offers a direct way of testing whether the restrictions imposed in equation (25) are, in fact, correct. Expression (24), multiplied by the number of families F, converges to a chi-square distribution with $S^2K-K(S+1)$ degrees of freedom.

The major advantage of the correlated-random-effects procedure is that we avoid the differencing inherent in the fixed-effects procedure and at the same time eliminate the bias arising from a family fixed effect.

V. Data Set and Descriptive Analysis

In this section, we describe the construction of our sample and the features of the data. The analysis is based on the *German Socio-Economic Panel* (GSOEP), a longitudinal study of private households in Germany. We construct two samples: A full sample that consists of persons between age 18 and 56, and a subsample that comprises only the subgroup of siblings.

1. Construction of the Working Samples

The structure of the GSOEP is very similar to the American *Panel Study of Income Dynamics* (PSID) (see Wagner et al. 1993 for a general description). The GSOEP started in 1984 in West Germany with 5,921 households, 12,290 interviewed persons over age 16, and more than 2,000 children under age 17 (GSOEP West). The West German sample also includes the subsamples of the five largest foreign nationalities living in Germany. In 1996, 4,445 households with 8,606 interviewed persons were left in the sample. In 1990, another 2,179 households with 4,453 persons from East Germany were added to the panel (GSOEP East). Moreover, two samples of German immigrants from the former Soviet Union were added in 1994 and 1995. Our analysis is based on the West German sample, since the educational system and the incentives to participate in schooling were very different in East Germany.

In order to clarify the construction of the siblings sample, the interview process of the GSOEP has to be considered. All household members who are at least 17 years old are interviewed individually. Children below age 17 belong to the sample and are interviewed as soon as they reach the age of 17. The GSOEP keeps track of all members of *original* sample households. If a member of an original household leaves the old household and forms a new one, the new household is added to the sample, including all new household members. The GSOEP has been particularly successful in following up the persons who formed new households, and we exploit this feature of the GSOEP to construct a siblings sample. We use the identification number of the biological mother to identify siblings. Thus, all persons in the sample who have the same mother are treated as siblings. Note that our sample also comprises children who were below age 17 when the panel started in 1984.

Over time, the children of the original sample households grow up and finish their formal education, eventually entering the labor force. As time goes by, an analysis based on siblings becomes feasible with reasonably large sample sizes. On the other hand, panel mortality reduces the number of respondents and aggravates sample selection. As a compromise, we have chosen the sample year 1996. We use all siblings in the data set and not just pairs with exactly two siblings. Thus, our siblings sample constitutes an *unbalanced* panel. Since the longitudinal dimension does not help to identify the parameters of interest, it suffices to use one wave.

Persons who do not report a positive wage are excluded from the working sample. Since apprentices earn a reduced salary in Germany while they receive formal training and on-the-job training, they are also excluded. Since early retirement often starts around age 57, we further restrict the sample to persons below age 57 in order to avoid a bias that may arise if retirement decisions depend on wages. The latter restriction is not binding in the siblings sample.

2. Variable Description

The dependent variable in our wage regressions is the hourly log wage (*logwage*) constructed from monthly gross income and the time actually worked per month.[12] The education variables are constructed from data on the individuals' education degrees. The degree is coded as the minimum number of years required for obtaining the degree and not with the time actually spent in an educational institution (*degree*). In the case of vocational training (e.g., apprenticeships), where about half of the time is spent working and not in training, we count only half of the time as education. An apprenticeship, for example, is counted as one-and-a-half years in the variable *degree* instead of three years. We also construct dummy variables indicating the highest formal degree.

[12] Returns to education based on hourly wages may be smaller than those based on monthly or annual earnings, if the better educated work longer.

Table 1: Variables used in the regressions.

Logwage	Log of hourly net wage
Degree	Years of formal education
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
Other variables	Interaction terms of *Female* and *Cohort* with *degree*
Parentsdegree	Mother's/father's years of education, analogous to *degree*, deviations from sample mean
Parentscollege	Dummy variable, 1 if father or mother have 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.

Since the different school types of secondary schooling are the main driving force of school quality, our data allow us to indirectly control for school quality. Thus, the duration of education captures more than just the quantity aspect. Secondary general school („Hauptschule”) is coded with 8 or 9 years, depending on the birth year, intermediate school („Realschule“) with 10, and grammar school (“Gymnasium“) with 13 years. A university degree is coded with 18 years (13 years plus 5), while a degree obtained from an applied college is coded with only 16 years (13 plus 3).

Table 1 shows the variables used in the empirical analysis. Since age and year of birth are not separately identified in a cross-section, we use only the birth cohort (*cohort*). This variable is thought to reflect different entry wages of cohorts. The effect of aging should be covered by the variable *experience*. We measure family background in two different ways: We construct variables similar to the *degree* variable for the parents' years of education, and we also use a dummy variable for the parents' college degree (*parentscollege*). These variables appear to capture the influence of family background reasonably well. Including other variables like the family size or the age difference between parents and children does not change the results significantly nor does it provide any further insights.

3. Descriptive Analysis

The working samples consist of 3,547 (*full sample*) and 496 persons (*siblings sample*), respectively. The size of the siblings sample is smaller for several reasons. First, it is a sample of relatively young persons. Second, many of the young persons did not finish their formal education by the year 1996. Third, the fixed-effects estimation requires at least sibling pairs with complete data. In the following, we describe the main characteristics of the two samples (see also table 2).

Table 2: Summary Statistics of the working samples.

	Full Sample (N=3547)				Siblings Sample (N=496)			
Variable	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
Wage	25.5	16.3	3	451	22.9	13.6	3	258
Degree	11.6	2.6	8	20	11.4	2.5	8	20
experience	21.8	13.0	0	55	14.3	8.2	0	41
Tenure	9.6	8.6	0	42	5.7	4.6	0	24
Cohort	58.7	9.8	40	78	67.2	4.8	49	77
Female	42%	0.49	0	1	32%	0.47	0	1
parentsdegree	9.1	2.5	0	18	9.2	1.8	3.5	18
parentscollege	7%	0.26	0	1	6%	0.25	0	1

A major difference between the two samples is the distribution of age. By construction, the siblings sample is much younger than the full sample, since the siblings sample is made up of children of panel households. Average ages are 29 years and 37 years, respectively. We observe only a few siblings above age 37. As a consequence of the difference in age distributions, the sample distributions of other variables differ as well. Mean experience (21.8 versus 14.3 years) and mean tenure (9.6 versus 5.7 years) differ markedly. The average duration of education in the full sample is slightly longer than in the siblings sample (11.6 versus 11.4 years). This is due to the fact that persons with higher education are slightly underrepresented in the siblings sample.

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. The means of gross monthly incomes are 4,234 DM in the full sample and 3,886 DM in the siblings sample. The level of wages in the siblings sample is 10 percent lower than in the full sample. This can be attributed to the shorter labor market experience of the persons in the siblings sample.

Family background variables are 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' average duration of education are 8.68 and 9.64 years, respectively. In the siblings 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. Since the siblings sample covers the cohorts born between 1949 and 1977, the sample persons have experienced the years after the big educational reforms.

The correlation between the education of parents and children is considerable as table 3 shows. A college degree earned by the father is still a good predictor for the schooling success of his children. The probability of a child reaching a college degree is 3.5 times higher if his or her father also has a college degree. The intergenerational correlation of secondary schooling is also strong. The odds of successful completion of grammar school (13 years) are 65% if the father has the same degree; the odds are only 13% if the father has only 9 years of schooling. Parental education may influence the schooling decision of children in several ways; one is the effect of parental education on family income. The effect of income on the education decision is hard to disentangle from the direct effect of parental education. Acemoglu and Pischke (2001) use variations in the American income distribution to identify the effect of income. Since the individual costs of higher education in Germany are lower than in the U.S., the pure income effect on the schooling decision is probably lower in Germany.

Table 3: Correlation between children's and parents' education in the full sample

Degree of …	Child	Father	Mother
Child	1.00		
Father	0.43	1.00	
Mother	0.40	0.69	1.00

Notes: Degree is measured as the minimum years of formal education required for highest degree

VI. Estimation Results

First, we discuss the regression results based on the full sample. These results show the relevance of family background and the interesting interplay between the level effects on wages and the marginal returns to education (slope effects on wages). Heterogeneity turns out to be an important feature of returns to education. Second, as a benchmark, we present the results based on the siblings sample using the same specifications as in the full sample. Third, the fixed-effects and the correlated random-effects estimates are provided. The marginal returns to education for all specifications are calculated from equation (7) using the point estimates of the coefficients and the sample means of the interaction variables. We also report the estimated standard errors of the marginal returns. The variables that interact with education (except for the dummy variables) are normalized around their sample means. Thus, the base case is a person at the sample mean whose dummy variables are zero. The estimated coefficient of education in the log wage equation directly yields this person's marginal return to education.

1. Results for the Full Sample

We first estimate a basic human capital earnings function similar to the standard Mincer equation (see equation (1) above). This simple specification assumes that marginal returns are homogenous. The results of the full sample are presented in the left panel of table 4. The coefficients are highly significant and have the expected signs. The estimated marginal return to education of 6.39 percent lies in the range of results reported in the German literature. There is a positive cohort effect, which indicates that younger cohorts earn higher wages, given experience and the other control variables.[13] The wage of women is estimated to be 16 percent lower than the wage of males. It is important to note that the women's lower *average* wage does not imply that their *marginal* return to education is lower. In order to uncover differences in marginal returns, it is necessary to explicitly allow for heterogeneous returns. This will be accomplished by introducing interaction terms with the education variable.

Table 4: Regression results for the Basic Mincer Model

Dependent Variable	**FULL SAMPLE**			**SIBLINGS SAMPLE**		
Logwage	coeff.	std.err.	t-val.	coeff.	std.err.	t-val.
Degree	.0639	.00264	24.20	.0456	.00811	5.62
Female	-.1620	.01401	-11.51	-.1355	.03540	-3.83
Experience	.0308	.00216	14.27	.0333	.00837	3.98
Experience²	-.0004	.00004	-9.03	-.0007	.00025	-2.65
Tenure	.0042	.00101	4.13	.0113	.00367	3.07
Cohort	.0044	.00119	3.68	.0024	.00617	0.39
Constant	1.9865	.04601	43.17	2.2098	.13809	16.00
	Number of obs. = 3547, R-sq. = 0.31, Root MSE = 0.375			Number of obs. = 496, R-sq. = 0.19, Root MSE = 0.362		

Notes: Regression with robust standard errors. Cohort is normalized to mean zero; see also table 1 for a description of variables.

[13] This is in line with the results of Fitzenberger, Hujer, MaCurdy, Schnabel (2001),who find that an additive cohort effect is operating in the West German wage distribution.

Table 5: Marginal returns to education for different specifications in the full sample

		Marginal Returns to Education
Mincer equation (1)		6.39% (0.26%)
with parents' years of education (2a)		6.15% (0.29%)
with parents' college dummy (2b)		6.28% (0.28%)
with parents' years of education and heterogeneity (interaction terms) (3a)	Males	5.66% (0.35%)
	Females	6.69% (0.50%)
with parents' college dummy and heterogeneity (interaction terms) (3b)	Males, parents without	6.11% (0.33%)
	with college	3,10% (0.88%)
	Females, parents without	7.14% (0.50%)
	with college	4.13% (0.92%)

Notes: Standard errors in parentheses. Marginal returns are evaluated at sample means of the relevant variables. In the case of dummies, the marginal returns are evaluated separately.

In table 5 we give an overview of the estimated marginal returns for the different specifications based on the full sample. Adding family background variables (highest degree of parents or college education) to the regressors does not change the estimated marginal return significantly (see specifications 2a and 2b in table 5).

Heterogeneous marginal returns can be modeled using interaction terms with education. We introduce interactions of education (*degree*) with:

- *gender* in order to allow for gender-specific marginal returns,
- *education* of parents (Parentsdegree, Parentscollege) in order to allow for an effect of family background on marginal returns, and
- *cohort* in order to allow for cohort-specific returns to education.

Table 5 shows that males and females differ in their respective marginal returns to education. The estimated returns for females are 1 percentage point higher than for males (in specification 3a as well as in 3b). Specification 3b makes visible the heterogeneity in marginal returns due to family background (see also table 6). The marginal return of children of college educated parents is 3 percentage points lower than that of children from other families. Females whose parents did not complete college have the highest return to education. Males whose father or mother completed college have a return of only 3.1 percent.

Table 6: Regression using PARENTSCOLLEGE and interaction terms with education.

Dependent Variable:	**FULL SAMPLE**		**SIBLINGS SAMPLE**	
Log wage	coefficient	t-value	coefficient	t-value
Degree	0.061	18.51	0.033	3.26
Female	-0.282	-4.26	-0.331	-2.21
Experience	0.030	13.47	0.034	4.17
Experience²	0.000	-8.37	-0.001	-2.56
Tenure	0.004	3.67	0.013	3.59
Cohort	0.017	5.06	0.048	2.26
Parentscollege	0.465	3.55	0.737	1.50
Heterogeneity terms:				
*Parentscollege *degree*	-0.030	-3.38	-0.033	-1.16
*Female * degree*	0.010	1.83	0.018	1.37
*Cohort * degree*	-0.001	-4.16	-0.004	-2.12
Constant	2.037	38.54	2.295	15.23
	Number of obs = 3547, R-sq. = 0.32, Root MSE = 0.373		Number of obs = 496, R-sq. = 0.23, Root MSE = 0.356	

Notes: Cohort is measured as deviation from sample mean. Regressions with robust standard errors.

In the left panel of table 6 we present the complete regression results for our preferred specification 3b, which includes the parents' college dummy and the corresponding interaction term.[14] The gender wage differential measured by the dummy *female* has increased to 28 percent compared to the basic Mincer specification. However, the interaction of the gender variable with education is positive (though at the margin of significance at the 5 percent level). An additional year of education has a higher pay-off for females than for males. This means that while females have lower wage levels than men, they have higher *marginal* returns to education. The gap in gender pay narrows with higher education. However, it does not close.

A similar interplay between level and slope effects is present in the case of birth cohorts. Younger cohorts have higher wage levels, but lower marginal returns. The estimated level of wages increases by 1.7 percent per birth year, while there is a counteracting negative effect of the birth year on the marginal return to education. Younger cohorts, who have acquired a

[14] The full results for the other specifications are available upon request.

higher level of education, could do so because of reductions in the private costs of education during the 1970s.[15]

Next, we consider the effects of parental education (i.e., family background) on wages and returns. First, the educational background has a direct effect on wages, as persons whose parents have a college degree earn significantly higher wages than others.[16] Note that this effect works in addition to the indirect effect of family background on wages through the level of education. Second, the return to education of persons with college-educated parents is considerably lower (about 3 percentage points). This result is robust across different specifications and is consistent with the literature using US data (Ashenfelter and Rouse 1998).

In the context of our theoretical model, the interplay between returns to education and levels of wages can be interpreted as follows: Returns to education are an increasing function of the marginal benefit parameter b_i and the marginal cost parameter r_i (see equation 7). Thus, the lower returns to education of children from well-educated families may be due to lower marginal costs *or* due to lower marginal benefits. The latter is highly implausible, since it would imply that children of better educated parents are less productive in school. Moreover, the optimal *level of schooling* depends positively on the marginal benefit parameter b_i, and *negatively* on the marginal cost parameter r_i (see equation 6). Since we observe a strong positive correlation between parents' and children's education, lower marginal returns can go along with higher levels of education only if marginal costs are lower for children from families with higher education. If, at the same time, those children have higher marginal benefit parameters than others, the marginal costs should be even smaller (see figure 1).

Given the egalitarian design of the German educational system, this is a very interesting and surprising result. Although there are no official tuition fees for higher education in Germany, and although there is an income support program for students, costs still seem to play an important role. There are several reasons that should be considered. First, financial restrictions are binding in many cases, since the income support program does not cover a large part of the lower middle class. Moreover, some programs, especially law schools,

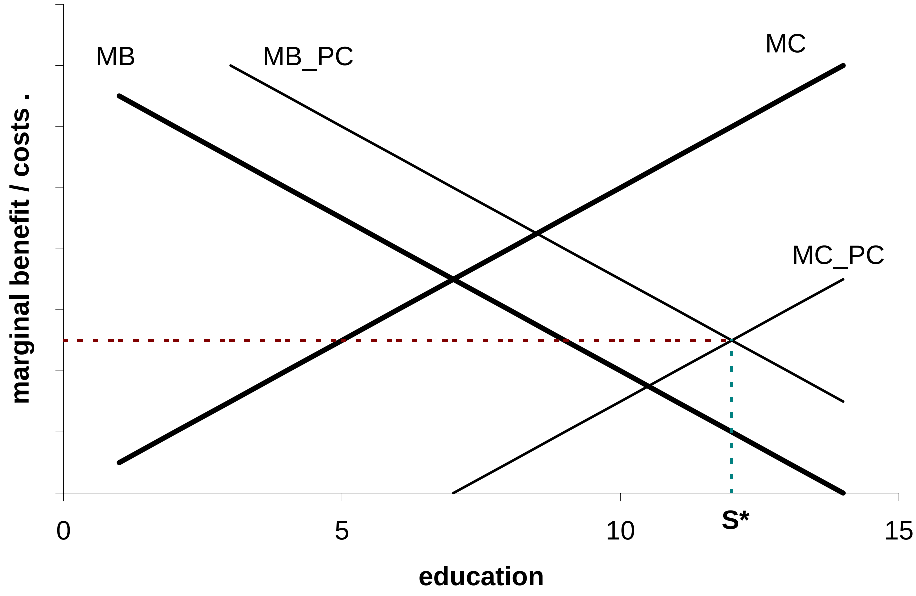
---

[15] Aggregate educational earnings differentials have been relatively stable over the last 25 years.

[16] The father's education has a significant positive effect on the *level* of wages while the mother's education is not significant. This result is also often reported in the literature using US data.

require additional private tuition in order to successfully complete the exams. Second, non-financial costs, such as taste factors and psychic costs, might play a very important role.[17]

Figure 1: Marginal returns, education decision and different family backgrounds



Notes: MB (marginal benefit function), MC (marginal cost function), MB_PC (marginal benefit function if Parents have College degree), MC_PC (marginal cost function if Parents have College degree), S* (optimal level of education for children if Parents have College degree).

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 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.

2. Results for the Siblings sample

Now we present the results for the siblings sample, which comprises 496 siblings from 224 different families (183 pairs, 34 triples, and 7 quadruples). We briefly discuss the results for the same specifications as in the full sample estimated by ordinary least squares (see above, tables 4 and 6). We do so in order to check how the results depend on the specific sample. Then, we present the fixed-effects and the correlated random-effects estimates for the siblings sample.

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

In the siblings sample, the estimates of the returns to education are lower than in the full sample. The coefficient of *DEGREE* is only 0.046 in the basic Mincer specification (table 4) and only 0.0319 in the specification with heterogeneity (table 6). The lower returns are due to the fact that the sample covers relatively young persons who earn less on average. The signs of the coefficients do not change, some coefficients rise in magnitude. However, the standard errors also increase due to the much smaller sample size of only 496 persons. The family background variables, as well as the heterogeneity terms, are jointly significant at the 5% level. Thus, the null hypothesis of homogeneous returns to education is also rejected in the siblings sample.[18]

The results differ between the full and the siblings sample for several reasons. First, the siblings sample is much younger, and, thus, experience and tenure are smaller. This may influence the estimates of returns to education if returns to *experience* depend on schooling. For instance, the wage-experience profile is steeper for college graduates (see Fitzenberger et al. 2001 for age-earnings profiles for German males). Second, the sample underrepresents persons with college degree. This effect is partly offset by a positive selection of persons who have completed their degree quickly. Third, the siblings 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 siblings sample has more weight (Hsiao 1986). This shifts the simple OLS-estimates towards the within-estimates. Given the high intra-family correlation, which is present in this siblings sample, the ordinary least squares estimator is clearly not appropriate. Thus, we now turn to estimators that explicitly deal with the family structure of the data set.

Next, we estimate the wage equation using a fixed-effects estimator for an unbalanced panel.[19] Family fixed effects absorb all family-specific differences between persons like parents’ educational background, family resources etc.; they also absorb individual-specific variation that is linearly related to family characteristics. Only heterogeneity within the family is left. As a matter of fact, we cannot estimate the impact of those components in a fixed-effects approach. Simply introducing **family fixed-effects** without any control variables „explains“ 41 percent of the total variance in log-wages; this underlines the importance of

[18] In specifications using other indicators for the parental education, the test yields similar results.

[19] Since a Hausman test on random-effects is rejected, we do not display the simple random-effects estimates here. They are of little interest given that they lie between OLS and Fixed-effects estimates, and given that we present *correlated* random-effects estimates below.

family background in determining wages. The fixed-effects results for the specification corresponding to table 6 are reported in table 7 below.

Table 7: Fixed-effects wage regression for siblings sample.

Dependent variable logwage	**Coeff.**	**Std. err.**	**t-val.**
Degree	.02423	.01648	1.470
Female	-.47194	.21152	-2.231
Experience	.03721	.00974	3.820
Experience²	-.00075	.00028	-2.710
Tenure	.01301	.00574	2.269
Cohort	.08097	.03153	2.568
Heterogeneity terms			
*Parentscollege * degree*	-.04207	.02766	-1.521
*Female * degree*	.02603	.01819	1.431
*Cohort * degree*	-.00596	.00254	-2.343
Constant	2.45318	.23222	10.564

Fixed-effects regression, 496 observations in 224 families, 2 to 4 siblings per familiy. R-sq. within = 17.5%, Wald test that family effects are zero (H_0: α_i=0): p-value = 2.93%.
Decomposition of residual variance: std.dev. of family effects = 0.345, std.dev. of idiosyncratic effects = 0.335, fraction of residual variance due to family effects = 51.5%.

The decomposition of the residual variance reveals that family-specific effects account for 51.5 percent of the residual variance of log wages. Thus, 48.5 percent of the residual variance are attributed to idiosyncratic effects, i.e., differences between siblings. The fixed-effects estimate for the coefficient of education (degree) is only 0.024 and is not significant at conventional levels. Recall that the corresponding OLS estimate was 0.033 in the siblings regression. This estimate of the marginal return to schooling corresponds to the base case (a male person with mean birth year, whose parents have no college degree). The effects of birth year on log wages and on the returns to education are similar to the estimates presented above: younger cohorts earn higher wages and reap lower returns to education at the margin.

Returns are still heterogeneous after controlling for family fixed-effects. The Wald test on homogeneous returns clearly rejects the null hypothesis at a significance level of 2.27 percent. In order to demonstrate the importance of heterogeneity, we again calculate the marginal returns for specific cases. For younger cohorts, the estimated returns are significantly lower. For a (mean age) son of college educated parents, the point estimate of marginal returns to education is –1.78% and is *significantly* different from zero. The estimated marginal returns to education for females are 5.03% or 0.82%, respectively, only the former being significantly different from zero.

The idiosyncratic effects are potentially correlated with education and may lead to an upward bias of the fixed-effects estimator. On the other hand, measurement errors in education tend to bias the estimate downwards. The net effect on the estimated marginal return is unclear; Card (1999) argues that the attenuation bias of measurement error is more severe and leads to a downward biased estimate. In order do deal with the problem of measurement error, we also estimate a correlated random-effects model.[20] The estimator is less responsive to measurement error and at the same time more efficient if the assumption of exogeneity of the control variables is met. Thus, we also expect to get more precise estimates.

Table 8: Correlated Random-Effects Specification for Siblings Sample.

Dependent variable logwage	**Coeff.**	**Std. err.**	**t-val.**
Degree	0,0279	0,0089	3,13
Female	-0,5060	0,1300	-3,89
Experience	0,0365	0,0068	5,40
Experience²	-0,0008	0,0002	-3,87
Tenure	0,0083	0,0037	2,21
Cohort	0,0291	0,0220	1,33
Parentscollege	0,2669	0,1879	1,42
Heterogeneity terms			
*Parentscollege * degree*	-0,0194	0,0139	-1,39
*Female * degree*	0,0274	0,0110	2,49
*Cohort * degree*	-0,0019	0,0018	-1,06
Constant	2,4465	0,1183	20,68

Notes: Correlated Random-Effects Estimation. 496 persons in 224 families. In addition to the above parameters, the correlation structure has been estimated, which entails 27 λ-parameters (see eq. 17). χ^2-Test of the overidentifying restrictions (see eqs. 25f): χ^2 = 97.93975 with ν=54, rejected at P-value = 0.082%. Wald-test on H_0 (homogeneous returns): W = 10.38 with ν=3, rejected at P-Value = 1.56%.

[20] There are other ways to mitigate the measurement error problem, e.g., using repeated observations.

Table 8 shows the results for the correlated random-effects estimation. The coefficients are estimated more precisely than in the fixed-effects case. The point estimate of the marginal returns is still low at 2.79 percent for the base case. While the coefficients of the gender variables have turned significant, the cohort effects are still individually insignificant. However, the family background variables are jointly significant and the homogeneity of returns is clearly rejected by a Wald test. The results, again, confirm the heterogeneity of returns to education.

Finally, we display the estimates of marginal returns for different estimators and samples in table 9. The upper panel in table 9 corresponds to the estimates reported in tables 6, 7, and 8. The lower panel additionally reports the estimates based on the degree of parents (*parentsdegree*) instead of the college variable. The latter specification yields similar qualitative results; however, it is better suited to visualize the heterogeneity due to family background (see also figure 2).

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

	FULL SAMPLE	SIBLINGS SAMPLE		
	OLS	OLS	Fixed Effects	Correlated Random Effects
Parentscollege				
+ Interaction Terms				
Males with				
Parentscollege = 0	6.11%	3.28%	2.42%	2.79%
	(0.33%)	(1.01%)	(1.65%)	(0.89%)
Parentscollege = 1	3.10%	-0.01%	-1.78%	0.85%
	(0.88%)	(2.75%)	(2.55%)	(1.23%)
Females with				
Parentscollege = 0	7.14%	5.06%	5.03%	5.52%
	(0.50%)	(1.36%)	(1.96%)	(1.19%)
Parentscollege = 1	4.13%	1.77%	0.82%	3.59%
	(0.92%)	(2.83%)	(2.71%)	(1.46%)
Parentsdegree				
+ Interaction Terms				
Males	5.66%	2.91%	2.72%	2.10%
	(0.35%)	(1.07%)	(1.56%)	(0.90%)
Females	6.69%	4.92%	5.66%	3.30%
	(0.50%)	(1.39%)	(1.90%)	(1.15%)

Notes: Standard errors in parenthesis. Continuous variables (Degree, cohort, parentsdegree) are evaluated at their sample means. Standard errors of marginal effects have been calculated using the estimated variance-covariance matrix of coefficients.

The most robust finding, which shows up in all variants of the model, is the considerable heterogeneity of returns to education, depending on factors such as family background, gender, and birth cohort. In all specifications, children from well-educated families display lower marginal returns to education than others. Moreover, females and older cohorts display higher marginal returns to education. In the context of our theoretical model, the low returns of children from well-educated families must be due to low marginal costs. As we have discussed before, a lower marginal benefit curve can be ruled out, especially because we observe that low marginal returns go along with high levels of education. Our results also imply that the low educational participation of children from working class households is mainly due to high marginal costs. If public policy intends to increase the participation of this group, it will be necessary to lower their marginal (financial and non-financial) costs.

The case of female participation in education is different, since the participation gap of former years has closed. The higher marginal return of females might be explained by higher marginal benefits; however, if this were the case they would display a higher educational participation than males. A more convincing explanation is positive self-selection in female labor participation: females with higher returns to education (i.e., the more productive) participate in the labor market with higher probability than those with lower returns. This self-selection effect is not captured in our analysis, since we do not model the participation decision. Clearly, this is an important step for future research.

Another important result concerns the magnitudes of different estimators, suggesting that the average marginal return to education may be overstated by OLS estimates of standard Mincer-type wage equations. Introducing family background and interaction terms to control for heterogeneous returns slightly reduces the estimates of returns to education. The low level of returns in the siblings sample is probably due to the particular sample, which over-samples young persons. Therefore, the siblings analysis underestimates returns to education. Yet, a comparison of OLS estimates in the siblings sample with other – presumably less biased – estimates from the same sample can give us an indication on whether the OLS estimates in the full sample including family background controls are still biased upwards. The fixed-effects estimation yields the lowest estimates of the returns to education of only 2.4% for the base case, and is barely significantly different from zero. The coefficient for the effect of parents' college is (implausibly) large and implies highly negative marginal returns to education for some groups. If parents' years of education are used instead of parents' college (see figure 2), the same extreme results appear in the fixed-effects estimation.

Figure 2: Marginal Returns to Education as a function of parents' years of education

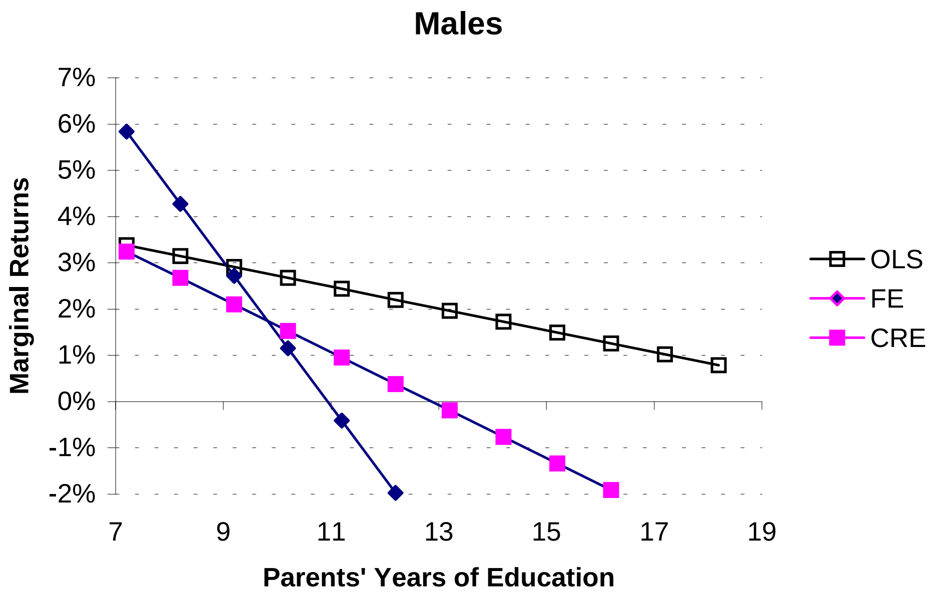
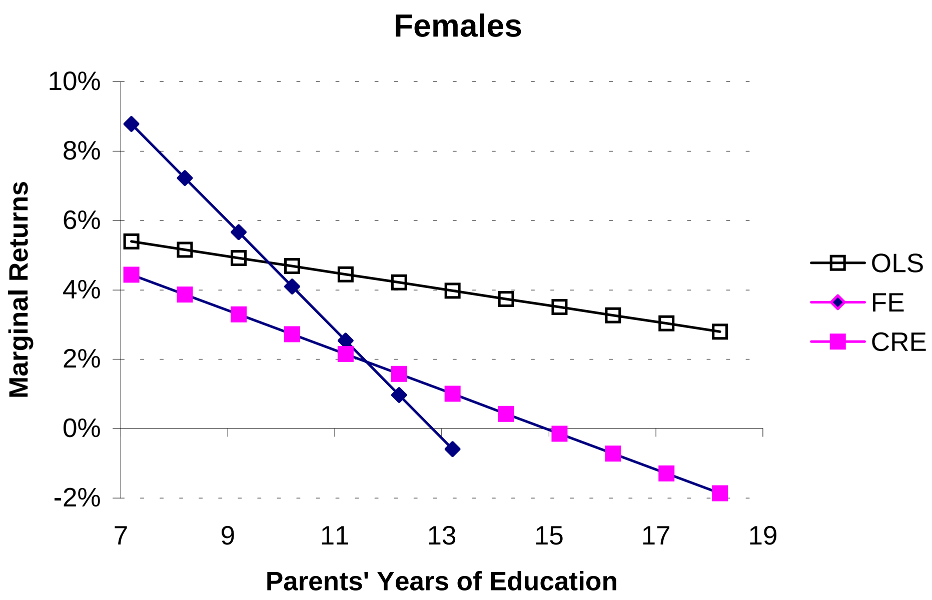


Figure 2 continued.



Notes: Marginal rates of return to education for three estimation methods in the siblings sample using parents years of education as control variable and interacted with degree of respondent. OLS (ordinary least squares), FE (fixed-effects) and CRE (correlated random-effects).

The correlated random-effects estimators yield more plausible results than the fixed effects estimator. In addition to methodological considerations, this is a further argument in favor of the correlated random-effects estimation. The correlated random-effects estimation yields magnitudes of coefficients that range between the two extremes of OLS and fixed-effects. This suggests that the OLS estimates from the full sample are biased upwards, too.

VII. Conclusions

With data from the GSOEP, we estimated returns to education using two different samples and a number of different estimation procedures. First, we estimated returns to education in a simple OLS framework based on the full sample. Second, we constructed a siblings sample, which allowed us to perform family fixed-effects and family correlated random-effects estimations to reduce the potential endogeneity bias inherent in OLS estimation. The results from our analysis confirm the presumption that the OLS estimates overstate true returns to education. The correlated random-effects approach appears to be the most appropriate, because measurement error seems to lead to a downward bias in the fixed-effects estimates.

Our main result is that family background still matters despite the attempts – or political claims – to equalize educational opportunities. Family characteristics constitute an important part of the variation in wages and in marginal returns to education, which confirms the important role of family background. Persons with well-educated parents tend to have lower returns to education and earn higher wages than persons with less-educated parents. Based on our theoretical model, we argued that this must be due to lower marginal costs of education in well-educated families. In addition, we found that gender matters for returns to education. While women, on average, earn lower wages than men, they have higher marginal returns to education. This may be explained by a self-selection of the more productive women into paid work – as opposed to household production. As a by-product of our analysis, the same interplay of wage levels and marginal returns was found for the effect of cohort membership on wages and marginal returns. The reduction in private marginal costs of education in the 1970s and 1980s has increased participation in schooling and thereby has reduced private returns. At the same time, macroeconomic growth has lead to an increase in the wage level of younger cohorts.

The reforms of the education system benefited above all the middle class. If public policy intends to increase the participation of children from poorer families, it will be necessary to lower their private marginal costs. Financial grants or loans may be one way to achieve that goal. The trend of restricting financial assistance to students is counterproductive as it leads to increased social selection and to an under-utilization of human resources.

VIII. References:

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

Acemoglu, D. and J.-S. Pischke, 2001, "Changes in the wage structure, family income, and children's education." *European Economic Review*, 45, 890-904.

Altonji, J. G. and T. A. Dunn, 1996, "The Effects of Family Characteristics on the Return to Education," *Review of Economics and Statistics*, 78, 692-704.

Ashenfelter, O. and A. Krueger, 1994, "Estimating the Returns to Schooling Using a New Sample of Twins," *American Economic Review*, 84, 1157-1173.

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

Becker, G. S., 1967, *Human Capital and the Personal Distribution of Income*. Ann Arbor, Michigan: University of Michigan Press.

Card, D., 1995, "Earnings, Education, and Ability Revisited," Research in Labor Economics, 14, 23-48, Solomon Polachek (ed.). Greenwich, Connecticut: JAI Press.

Card, D., 1999, "The Causal Effect of Education on Earnings," *Handbook of Labor Economics,* Orley Ashenfelter and David Card (eds.), Volume 3, Chapter 30, 1801-1863.

Card, D., 2001, "Estimating the Returns to Schooling: Progress on Some Persistent Econometric Problems," *Econometrica*, 69(5), 1127-60.

Fitzenberger, B., R. Hujer, Th. MaCurdy, and R. Schnabel, 2001, "Testing for Uniform Wage Trends in West-Germany – A Cohort Analysis using Quantile Regressions for Censored Data," *Empirical Economics*, 26 (1), 41-86.

Fitzenberger, B., R. Schnabel, and G. Wunderlich, 2001, "The Gender Gap in Labor Market Participation and Employment: A Cohort Analysis for West Germany," *ZEW Discussion Paper* No. 01-47.

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

Hsiao, C., 1986, *Analysis of Panel Data.* Cambridge: Cambridge University Press.

Ichino, A. and R. Winter-Ebmer, 1999, "Lower and upper bounds of returns to schooling: An exercise in IV estimation with different instruments." *European Economic Review*, 43, 889-901.

IW – Institut der deutschen Wirtschaft, 2002, *Deutschland in Zahlen.* Köln: Deutscher Instituts-Verlag.

Lauer, C. and V. Steiner, 2000, "Returns to Education in West Germany – An Empirical Assessment." *ZEW Discussion Paper* No. 00-04.

Licht, G., and V. 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, J., 1974, *Education, Experience and Earnings*. New York: Columbia University Press.

Nickell, S. and B. Bell, 1996, “Changes in the Distribution of Wages and Unemployment in OECD Countries,” *American Economic Review Papers and Proceedings*, 86(2), 302-308.

Schimpl-Neimanns, B., 2000, “Hat die Bildungsexpansion zum Abbau der sozialen Ungleichheit in der Bildungsbeteiligung geführt?”, ZUMA-Arbeitsbericht 2000/02, Mannheim.

Wagner, G., R. Burkhauser, and F. Behringer, 1993, “The English Language Public Use File of the German Socio-Economic Panel Study, ” *The Journal of Human Resources*, 28(2), 429-433.

Wissenschaftsrat, 2002, *Eckdaten und Kennzahlen zur Lage der Hochschulen von 1980-2000*, Köln.