

Blaess, Virginie

Working Paper

Siblings and Educational Attainment in West Germany

Discussion Paper, No. 2005,001E

Provided in Cooperation with:

University of Erfurt, Faculty of Economics, Law and Social Sciences

Suggested Citation: Blaess, Virginie (2005) : Siblings and Educational Attainment in West Germany, Discussion Paper, No. 2005,001E, Universität Erfurt, Staatswissenschaftliche Fakultät, Erfurt

This Version is available at:

<https://hdl.handle.net/10419/23938>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



UNIVERSITÄT
ERFURT

Staatswissenschaftliche Fakultät

Faculty of Economics,
Law and Social Sciences

Discussion Paper No.: 2005-001E

**Siblings and Educational Attainment
in West Germany**

Virginie Blaess

Für den Inhalt des Diskussionspapiers sind die jeweiligen Autoren/innen allein verantwortlich.

ISSN 1610-9198 (Print)

ISSN 1610-918X (Internet)

Siblings and Educational Attainment in West Germany

Virginie Blaess ^{*†}

February 15, 2005

Keywords: Unobservable heterogeneity, matching, ordered-probit
JEL classification: C13, C14, I20

Abstract

Individual decisions on education are still an important topic in social sciences research. Our goal is an analysis of the impact of siblings on educational attainment in West Germany. Theories of educational decisions in a family context suggest several possible effects of siblings. During the 1990s, several authors analyzed this relationship for the United States and came up with contradictory results on the relative importance of different factors. Consequently, an empirical analysis is required, which is provided in this paper based on data from the GSOEP. In order to control for unobservable heterogeneity in educational decisions, several empirical specifications including propensity score matching are tested. The results suggest that boys are favored by their parents relative to girls. Furthermore, the gender of their siblings shows no significant impact on the educational attainment of boys, while a significant effect is found for girls. Finally, the educational attainment of an elder sibling shows a significant and positive effect on education decisions of the second child.

^{*}Department of Economics, Law, and Social Sciences, University of Erfurt, P.O. Box 900 221, 99105 Erfurt, Germany, Phone: +49/361/737-4598, E-Mail: virginie.blaess@uni-erfurt.de

[†]I would like to thank the participants of the Sixth Conference on Education, the participants of the Sixth GSOEP Users Conference and the participants of various seminars in Strasbourg, Osnabrück and Erfurt for valuable comments and suggestions. Furthermore, I particularly thank François Laisney and Peter Winker for fruitful suggestions which helped to improve my paper.

1 Introduction

Many authors, from various fields of social sciences, deal with the analysis of the relation between sibling composition and educational attainment. In fact, this relationship has been studied in economics, sociology and psychology, and this for several decades. For example, Powell and Steelman (1989) refer to the writings of Dumont and Galton at the end of the eighteenth century. Given the abundance of the different works completed in this fields, various theories have been proposed. All these theories converge in the statement that family size has a negative impact on the educational attainment of the children.¹

The sociolocal literature puts forward the resource dilution model (Powell and Steelman, 1990). This model affirms that educational attainment of children hinges on inputs of time and money from parents. Therefore a negative relation between the size of the sibship and the educational attainment of the children should be expected.

In economics, Becker (1981) and Becker and Tomes (1976) present the parental allocation of time model. Parents care about the number and the ‘quality’ of the children. Given the budget constraint $p_c q n + \pi_z Z = I$,² the family maximizes an utility function $U = U(n, q, Z)$. The result produces a compromise between the number of children and the ‘quality’ of children in a family. In order to maximize this utility, three types of behavior are observed: i) a non-discriminatory time allocation attitude: independently from the child’s results, each child becomes the same quantity of time, ii) an achievement maximization attitude: there is a positive dependency between the productivity of the child and the time distribution, iii) a compensatory resource allocation attitude: the child with the greatest difficulties obtains the most attention.

The psychological literature describes the confluence model.³ This model evaluates the family’s intellectual environment. This environment is equivalent to the weighted average of the mental ages of all members of the family. Thus, there is a negative relationship between the number of children and the average mental age of the family. This average mental age determines the intellectual development of the children. If the average mental age of the

¹For more information, see Iacovou (2001) or Conley (2000).

² p_c is the cost of an unit of quality, n the quantity of children, q the quality of children and Z the quantities of other commodities.

³For further information, see Powell and Steelman (1990).

family is low, the intellectual development of the children will be also weak. Powell and Steelman (1989) bring out another theory: the normative climate hypothesis. The position allocated to a child in the family depends on her sex, on the birth order and on the number of siblings. Finally Rosenberg's sex minority hypothesis (1965) says that if an individual, given her sex, is a minority in her sibling composition, then she will receive a special attention. All models suggest that children's attainment declines as the number of sibling increases, the increase in the number of children reduces the resources granted to each child. It should however be noted that the influence exerted by a sister or a brother can be positive, *ceteris paribus*. Iacovou (2001) alludes to the fact that educational attainment of only child is lower.

A great number of empirical studies which present this relation for the US is available,⁴ but only few studies were carried out in this field for Germany.⁵ The purpose of this paper is the analysis of exactly this relation (between educational attainment and the number and gender of siblings) in West Germany, using data from the German Socio-Economic Panel (GSOEP). Only a limited number of background and individual characteristics is observable, i.e. family social class, educational attainment of parents, sex, age of the respondent. Parents' background characteristics also influence the educational attainment of children. Their educational attainment or the income of the family are significant factors. Nevertheless, some of the parental background characteristics which influence the sibling composition are unobservable.⁶ Moreover, resources granted to each child depend not only on the sibling composition but also on the preferences which parents have for a child. Unfortunately these preferences are not observable. Thus the study of the relationship between educational attainment and sibling composition implies the analysis of parental and individual characteristics. This analysis is to some extent limited as many characteristics are unobservable. Besides the problem of unobservability, characteristics also differ between individuals, i.e. determinants of the sibling composition, preferences of individuals or scholar abilities. To control for this resulting unobservable heterogeneity, a non-parametric method is adopted in this paper, i.e. the matching procedure, which has not been used in this context before.

The structure of this study is as follows: Earlier empirical works are

⁴Butcher and Case (1994), Conley (2000), Hauser and Kuo (1995), Kaestner (1996).

⁵Bauer and Gang (2001).

⁶For example the fertility of parents.

presented in Section 2. Section 3 describes the methodological framework. Data and results are described in Section 4. Section 5 presents a short analysis of the effect of an elder sibling on educational attainment of a child and Section 6 concludes.

2 Some empirical works

Various research carried out for the US are presented here. Butcher and Case (1994) study the impact of the siblings composition on the educational attainment of men and women in the United States. They base their study on models of parental preferences or of strategies for parental investment in children's education.⁷ If the family faces borrowing constraints, investments in children's education are limited, therefore families tend to favor children with the highest marginal return to education. If parents face no borrowing constraints, two types of behavior can be observed. Educational investment in this case will be made until the marginal return to education for each child is equal to the market rate of interest. However, parents will tend to favor the weakest child if they want to reduce inequalities between children. For their study, Butcher and Case use the data from the Panel Study of Income Dynamics (PSID),⁸ the National Longitudinal Survey of Women (NLSW)⁹ and the Current Population Survey (CPS). They run regressions on educational attainment in order to display the effects of the number of siblings, the effects of indicators for any sisters or for any brothers and finally the effect of the percentage of sisters in the sibship. Furthermore, they run logistic regression analyses of three educational transitions.¹⁰ Finally, Butcher and Case study the impact of the siblings composition on other socioeconomic outcomes, e.g. earnings and working situation. Above all, they observe an increase of individual educational attainment over the last seven decades. Additionally, parental's schooling positively influences educational attainment of their children. Furthermore, they find that 'throughout the century women's educational choices have been systematically affected by

⁷Models presented by Becker (1981).

⁸In this panel, Butcher and Case (1994) observe women and men aged between 24 and 65 in 1985.

⁹In this survey they observe women aged between 30 and 44 in 1967.

¹⁰The transition to high school graduation, college attendance and college graduation.

the sex composition of their siblings, and that men's choice have not'.¹¹

Hauser and Kuo (1995) replicate the regressions conducted by Butcher and Case (1994) using different data. They use the Occupational Changes in a Generation Survey (OCG), the 1986 to 1988 Surveys of Income and Program Participation (SIPP) and the 1989 National Survey of Families and Household (NSFH). They cannot confirm the results obtained by Butcher and Case. Instead, their results show a composition of positive and negative effects of the sibling sex composition on educational attainment. There is no evidence that the presence of sisters affect women's or men's educational attainment.

Kaestner (1996) resumes the work of Butcher and Case with more recent data from the National Longitudinal Survey of Youth.¹² Besides, he studies the effect of the sibling sex composition on educational attainment for children, teenagers and black individuals. Finally he examines the relation between sibling sex composition and family resources, and between birth spacing and sibling sex composition. He cannot support the results provided by Butcher and Case, as he finds no effect of sibling composition and educational attainment on white individuals. He finds that the presence of sisters has a significant positive effect on the educational attainment of black teens.

Conley (2000) discusses the three previous studies and tries to expand the analysis. In order to conduct his analysis, he uses respondents of the 1989 wave of the Panel Study of Income Dynamics (PSID). Just like Kaestner and Hauser and Kuo, he does not confirm the results from Butcher and Case. Instead, he reports that it is the number of opposite sex siblings that has the strongest effect on educational attainment. His results suggest that the presence of sisters (brothers) in the sibship disadvantages men's (women's) educational attainment. That is, 'gender effects manifest only in relation to the sex of the respondent herself'.¹³

Finally for Germany, Bauer and Gang (2001) conduct the same type of analysis. They compare the effects of sibling composition on educational attainment between 'cultural affiliation' (West German, East German, and foreign origin).¹⁴ Using data from the GSOEP, they conduct their study with individuals aged between 17 and 46. Their results show that the 'cultural affiliation' influences the relation between sibling sex composition and edu-

¹¹Butcher and Case (1994), p.531.

¹²Youth born between the years 1957 and 1965.

¹³See Conley (2000), p. 454.

¹⁴See Bauer and Gang (2001), p.1.

cational attainment. However, their results expose that only the presence of sisters in households with more than two children has a significant negative effect on West German males, and a significant positive effect on foreign females. They find no significant relation between sibling sex composition and educational attainment in Germany.

3 Methodological framework

For comparison with the previous empirical findings reported above, a standard linear regression model is used to estimate effects of siblings composition (i.e. number of sisters, of brothers, the presence of sisters, of brothers, percentage of sisters) on educational attainment (educational attainment is determined in years: 7 years correspond to the lowest level and 18 years correspond to an university degree) given social backgrounds (i.e. schooling duration of the parents, occupation status of the father), and other control variables (i.e. sex, age of the respondent). However, in contrast to earlier studies, the likely presence of unobserved heterogeneity in the sample, which might be correlated both with the decision about the children and the investment in their education, is take to account. To this end, a matching procedure is used. Finally, to analyze the effects of an elder sibling on educational attainment of a child, an ordered probit model is applied.

3.1 Matching

In order to estimate the effects of family size and family composition on educational attainment it is necessary to take account of parental and individual characteristics (such as preferences and abilities). Most of these characteristics are unobservable, and differ across individuals. It is thus necessary to take this unobserved heterogeneity into consideration. The matching method permits us to control it. This matching procedure is a non-parametric method (Blundell, Dearden and Sianesi, 2001, and 2003), which enables us to cope with the unavailable information. The method is based on the model of causal effects or of potential outcomes from Roy (1951) and Rubin (1974).

3.1.1 The model of causal effects

Two potential outcomes are considered: Y_i^1 , if there is participation in a treatment (in our case to have at least a sister or a brother), and Y_i^0 , if there

is no participation. Participation is described by the variable:

$$S_i = \begin{cases} 1 & \text{if participation} \\ 0 & \text{otherwise.} \end{cases}$$

The treatment or causal effect for i is:

$$Y_i^1 - Y_i^0. \tag{1}$$

Given X_i the set of observable characteristics of i , the average treatment effect is:

$$E(Y_i^1 - Y_i^0 | X_i), \tag{2}$$

and the average treatment effect on the treated is given by:

$$E(Y_i^1 - Y_i^0 | X_i, S_i = 1). \tag{3}$$

3.1.2 The problem of identification

In order to identify the average causal treatment effect,¹⁵ Lechner (2002a) considers identification under the conditional independence assumption, given by:

$$Y^0, Y^1 \perp S | X = x, \forall x \in \chi. \tag{4}$$

X is a vector of characteristics in a particular characteristic space χ . This assumption asserts that the treatment participation and treatment outcome are independent conditional on a set of observable characteristics. Conditional on observed attributes X , the distribution of the outcomes in the treated group is the same as the distribution of the outcome in the non-treated group.

Heckman, Ichimura and Todd (1998) propose a weaker version of this assumption:

$$E(Y^l | X, S = 1) = E(Y^l | X, S = 0) \quad l = 0, 1. \tag{5}$$

The next assumption:

$$P(S = 1 | X = x) < 1, \forall x \in \chi \tag{6}$$

ensures, given the distribution of characteristics in the treated sample, the existence of a control group.

¹⁵This treatment is generally not identified, only one of the potential outcomes exists.

3.1.3 The propensity score

Rosenbaum and Rubin (1983) introduce the concept of propensity score in order to reduce the dimensionality of the estimation problem. The propensity score is the probability that a person undergoes the treatment:

$$p(x) = P(S = 1|X = x) \quad (7)$$

The propensity score is a particularly important balancing score.¹⁶ Rosenbaum and Rubin showed that if the two outcomes are independent conditional on X (the conditional independence assumption is valid), the treatments are independent conditional on specific functions of X (the balancing score).

Under assumptions (4),(6) and (7) the propensity score matching is given by:

$$Y^0, Y^1 \perp S | p(x), \forall x \in \chi. \quad (8)$$

The conditional independence assumption is thus valid if the propensity score ($p(x)$) is used for matching rather than the complete vector of characteristics X .

3.1.4 The estimators

The matching procedure consists in comparing each treated individual i with some group of non-treated individuals j , whose characteristics are so similar as possible to the characteristics of the treated individual. It is then possible to compare the outcome Y_i of i with a matched outcome \hat{Y}_i given by the outcomes of the group of non-treated individuals. For the treatment, we define a binary indicator variable ‘any sisters’ (‘any brothers’) representing the presence of at least a sister (a brother) in the family, such that the treated are individuals who have at least a sister (a brother).

The estimator is given by:

$$\hat{\beta}_M = \frac{1}{\#(S_i = 1)} \sum_{i \in \{S_i = 1\}} \{Y_i - \hat{Y}_i\}, \quad (9)$$

where the matched outcome is given by:

$$\hat{Y}_i = \sum_{j \in C^0(p_i)} W_{ij} Y_j, \quad (10)$$

¹⁶A balancing score is a function of the characteristics.

where $C^0(p_i)$ is the set of neighbors of treated i in the control group, and $W_{ij} \in [0, 1]$ is the weight on individual j in forming a comparison with treated i and that:

$$\sum_{j \in C^0(p_i)} W_{ij} = 1.$$

The weight allocated to the control group depends on the used estimators.

The traditional matching estimator is the nearest-neighbor matching estimator.¹⁷ This estimator associates to the outcome Y_i of treated i a matched outcome given by the outcome of the most observably similar control unit k_i .

$$C^0(p_i) = \{k_i \in (S = 0) : |p_i - p_j| = \min_{k_i \in \{S=0\}} |p_i - p_{k_i}|\} \quad (11)$$

with $W_{ij} = 1$ if $(j = k_i)$. To exclude observations, which are not similar enough, an alternative to the nearest-neighbor matching estimator is the caliper matching estimator. This estimator ensures that the differences between the observations is not greater than the caliper. For a pre-specified caliper $\delta > 0$,

$$\delta > |p_i - p_j| = \min_{k_i \in \{S=0\}} |p_i - p_{k_i}|. \quad (12)$$

Another type of matching estimators are the kernel based matching estimators (Heckman et al., 1998). The idea is to associate to the outcome Y_i of i a matched outcome given by a kernel-weighted average of the outcomes of all non-treated individuals, where the weight given to non-treated j is in proportion to the resemblance between i and j :

$$W_{ij} = K((p_i - p_j)/h) / \sum_{j \in C^0(p_i)} K((p_i - p_j)/h), \quad (13)$$

where h is the bandwidth and $K()$ is the Kernel. The kernel functions most often used are the Gaussian kernel and the Epanechnikov kernel. The Gaussian kernel or normal kernel is given by: $K(u) \propto \exp(-u^2/2)$. All the non-treated individuals are used. The Epanechnikov kernel is given by: $K(u) \propto (1 - u^2) \cdot 1(|u| < 1)$. Only those non-treated units whose propensity score falls within a fixed caliper of h from p_i are used, $C^0(p_i) = \{j \in S_j = 0 : |p_i - p_j| < h\}$.

¹⁷For further information, see Blundell et al. (2001).

3.2 Ordered Probit

The last step consists in the analysis of the effect of the educational attainment of an elder sibling on the educational attainment of a child. The idea is to model a discrete choice among a set of alternatives. Each individual chooses, given family constraints, sex constraints and age constraints, the educational level she wants to reach. Educational attainment is the dependent variable. It is possible to order this choice variable. The variable education resumes these ordered choices. Five levels are defined.¹⁸ The first level corresponds to the ‘Hauptschulabschluss’ (secondary school certificate) or less (no vocational degree), the second level refers to the ‘Realschulabschluss’ (junior high school certificate) or the apprenticeship level, the third is associated with the ‘Abitur’ (general qualification for university entrance), the fourth level refers to the ‘Fachhochschulabschluss’ (advanced technical college certificate), and the last level is associated with an ‘Universitätsabschluss’ (university certificate).

We estimate relationships between this ordinal dependent variable (education) and a set of independent variables (family characteristics like parents’ years of schooling, occupation status of the father, eldest’s years of schooling, sex, age). In an ordered probit, the unknown cut-points k_j are estimated together with the parameter vector β :

$$y_i^* = \beta' x_i + u_i, \quad (14)$$

where y^* is unobservable. If the responses are coded 0,1,2,3,4, the observed variable becomes:

$$y = \begin{cases} 0 & \text{if } y^* \leq 0 \\ 1 & \text{if } 0 < y^* \leq k_1 \\ 2 & \text{if } k_1 < y^* \leq k_2 \\ 3 & \text{if } k_2 < y^* \leq k_3 \\ 4 & \text{if } y^* > k_3. \end{cases} \quad (15)$$

y^* is identified up to the cut-points k_j , because y^* depends on the established categories or levels j . In the case of an ordered probit, with $Var(u) = 1$ and given Y_i the dependent variable for individual i , the probability of observing an observation is:

¹⁸This classification follows Lauer (2002).

$$\begin{aligned} Pr(Y_i = y_j) &= Pr(k_{j-1} < \beta' x_i \leq k_j) \\ &= \Phi(k_j - \beta' x_i) - \Phi(k_{j-1} - \beta' x_i), \end{aligned} \quad (16)$$

where $\Phi()$ is the standard normal cumulative distribution function.

I_{ij} is an indicator variable:

$$I_{ij} = \begin{cases} 1 & \text{if } [Y_i = y_j]; \\ 0 & \text{otherwise,} \end{cases} \quad (17)$$

the log-likelihood function¹⁹ for the model is:

$$\ln L = \sum_{i=1}^n \sum_{j=1}^J I_{ij} \ln[\Phi(k_j - \beta' x_i) - \Phi(k_{j-1} - \beta' x_i)]. \quad (18)$$

4 Data and Results

4.1 Descriptive Statistics

The data set is drawn from the GSOEP (German Socio Economic Panel). The GSOEP was started in 1984 as a longitudinal survey of private households and persons in Germany. The central aim of this panel study is to collect representative micro-data on persons, households and families in order to measure stability and change in living conditions.

The sample is constructed from the 13th wave of the GSOEP (1996). This wave is interesting because it provides information about siblings. There is no direct information on parental income, and on the mother's professional status when respondents were 14. Therefore, only the father's occupational status is used as indicator of the financial standing of the family.

The sample is restricted to individuals of West German nationality, aged between 22 and 56, who finished their studies and work. Individuals who resume their studies after 1996 are excluded from the sample. The full sample consists of 3850 persons. Summary statistics are presented in the Tables 1 and 2.

¹⁹For further information, see Maddala (1999).

Table 1: Summary statistics

Variable^a	Mean	St.D.
Control Variables		
School	12.041	2.506
Sex	0.431	0.495
Age	37.849	9.161
Cohort 1	0.240	0.427
Cohort 2	0.345	0.475
Cohort 3	0.415	0.493
Education	2.638	1.049
The family		
Any sisters	0.553	0.497
Any brothers	0.574	0.494
Number of sisters	0.840	1.104
Number of brothers	0.877	1.105
Number of siblings	1.718	1.757
Percentage of sisters	47.467	42.215
Education of the respondent		
No degree	0.018	0.135
Low school degree	0.401	0.490
Apprenticeship	0.520	0.500
Technical college	0.239	0.427
Interschool	0.346	0.476
Professional college	0.059	0.236
High school	0.175	0.380
High Technical degree	0.059	0.236
University degree	0.100	0.300

to be continued...

^aThe definition of variables used in the regressions is provided in Table 10 in the Appendix.

Table 1 continued...

Variable	Description	Range
Education of the mother		
No info	0.065	0.247
University degree	0.012	0.107
Educational attainment	9.995	1.780
Education of the father		
No info	0.088	0.283
University degree	0.043	0.203
Educational attainment	10.754	2.288
Profession of the father		
Missing	0.250	0.283
Farmer	0.030	0.170
Worker	0.360	0.480
Employee	0.040	0.210
Middle position	0.150	0.350
Superior position	0.100	0.290
Self-employed	0.070	0.260
N	3850	

Figure 1 shows educational attainment by cohort and by sex. From this figure, it is possible to observe an increase of individual educational attainment in the last six decades in West Germany. However, a reduction is observable since the 1980s. In fact, mean years of educational attainment of individuals born before 1951 is 11.84; 12.18 for individuals born between 1951 and 1962; 12.03 for individuals born between 1962 and 1972. The school attainment of women has increased distinctly, the average school attainment grew from 11.32 years for women born before 1951 to 12.04 years for women born after 1962. Such an increase is not observable for men, since the average school attainment grew only from 12.12 to 12.34 for the first two cohorts, finally dropping and reaching 12.03 years for men born after 1962. Women's education is lower than men's for the first two cohorts, but reaches the men's level for the last cohort. One reason is that since the end of the forties, the share of women in students increase. Indeed, statistics observed since the fifties in West Germany show that the share of women obtaining a diploma is 25% in 1950, and is 43% in the eighties.²⁰ Moreover, this evolution also

²⁰Statistics presented by the 'Bundesinstitut für Berufsbildung' (Jan. 2004).

corresponds to the period when women increasingly participated in the labor market.

Figure 1: Educational Attainment by Cohort and by Sex

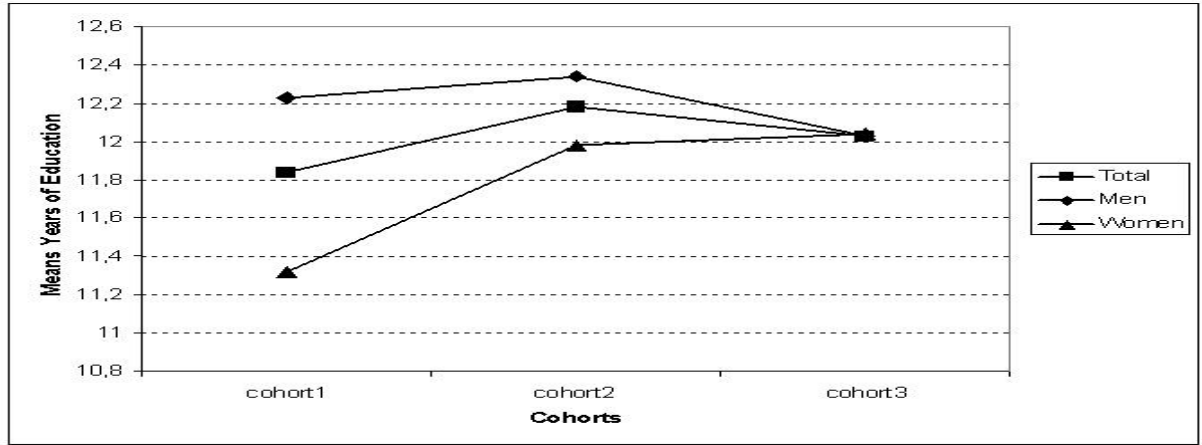


Table 2: Summary statistics: Men and Women

Variable	Mean	St.D.	Mean	St.D.
	Men		Women	
Educational attainment (in Years)	12.186	2.586	11.850	2.382
Age	37.957	9.070	37.708	9.281
Any sisters	0.560	0.497	0.545	0.498
Any brothers	0.572	0.495	0.578	0.494
Number of sisters	0.858	1.111	0.817	1.095
Number of brothers	0.886	1.125	0.866	1.079
Number of siblings	1.744	1.802	1.683	1.697
Percent of sisters*	37.490	40.298	60.631	41.074
Education	2.727	1.076	2.521	1.000
N	2190		1660	

*Percentage of sisters in the sibling including respondent.

Figure 2 shows educational attainment by the number of siblings in the family. One observes on this Figure an inverse relationship between education attainment and the number of siblings when there are more than two children in the family. One observes first an increase from the educational

attainment from one-child families to two-child families, and then a decrease. The level of education of the women remains lower than that of the men, but the same type of relationship between educational attainment and number of siblings is observable for both. The relation between siblings and educational attainment is presented in Table 3, one can observe that the number of sisters has a positive influence, with a fixed number of brothers and sisters (except for 3 sisters). Nevertheless, this Table presents no significant differences between the sibling composition and the educational attainment. The missing significance can be explained by the small size of the samples.

Figure 2: Educational Attainment by Number of Siblings

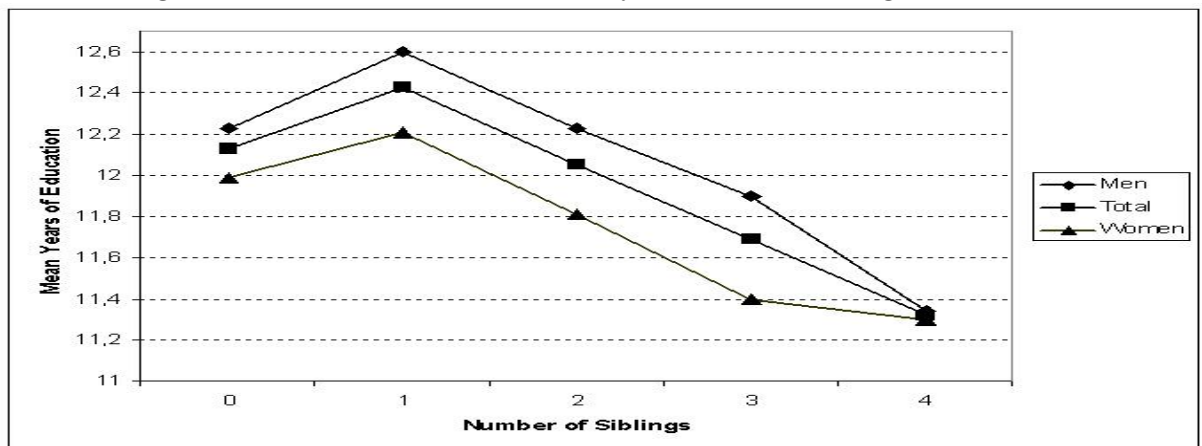
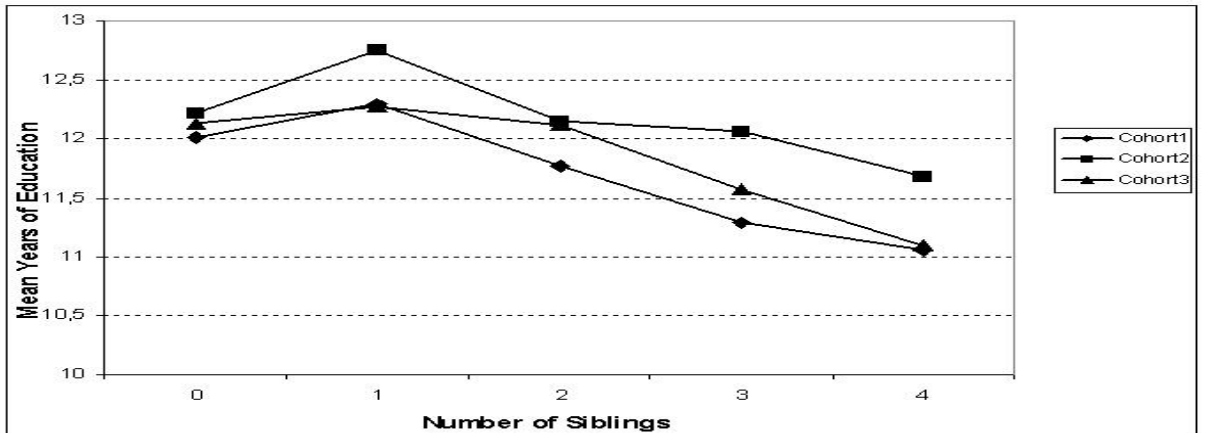


Table 3: Relationship between siblings and educational attainment

All the Sample				
Number of siblings	Gender sib.	Mean	(St.D.)	Obs.
0 sibling		12.13	(2.50)	915
1 sibling		12.43	(2.60)	1206
	0 sist.	12.39	(2.54)	624
	1 sist.	12.48	(2.60)	582
2 siblings		12.05	(2.57)	882
	0 sist.	11.70	(2.43)	213
	1 sist.	12.17	(2.59)	491
	2 sist.	12.18	(2.66)	178
3 siblings		11.69	(2.30)	386
	0 sist.	11.16	(1.82)	51
	1 sist.	11.82	(2.41)	136
	2 sist.	11.90	(2.52)	151
	3 sist.	11.23	(1.49)	48

Figure 3 shows educational attainment by number of siblings and by cohort. One observes on this figure that educational attainment increases from the first to the second cohort, and decreases afterwards. Nevertheless, the relationship between educational attainment and number of siblings remains the same for the three cohorts.

Figure 3: Educational Attainment by Number of Siblings and by Cohort



4.2 Results

4.2.1 OLS Estimations

As a first step of the empirical analysis, a standard linear regression model is estimated as in previous empirical work. For this analysis, families with an only child are excluded from the sample. Remaining the regressions from Conley (2000) on German data provide the results depicted in Table 4.

Table 4: Effects of the sibling composition on educational attainment

Column Variable	(1) Coef.	(2) Coef.
Age	0.005	0.005
Father's Education	0.200***	0.202***
Mother's Education	0.195***	0.195***
Sex	-0.331***	-0.328***
Number siblings	-0.162***	-0.248***
Number sisters		0.171***
Constant	8.676***	8.666***
R^2	0.233	0.236
N	2935	2935

The occupational status of the father is also used in the regressions.

Significance Level: * = $p < 10\%$, ** = $p < 5\%$, *** = $p < 1\%$.

Like Conley's results, the findings for Germany in column (1) of Table 4 indicate a significant sibling size effect on educational attainment. Each additional sibling costs 0.16 of a year of education, all else equal.

Opposed to Conley's results, column (2) shows that the number of sisters in the family is significant and positive, supporting the human capital model presented by Becker (1964). Additional brothers should disadvantage the individual's educational attainment.²¹ One potential explanation is that parents tend to favor the education of the boys.

Next a regression using dummy variables presented by Butcher and Case (1994) is run, i.e. any sisters, any brothers. The results of this regression are presented in Table 5. These variables (any sisters and any brothers) are significant, but when running these regressions separately for men and women, the variables remain significant for women, but not for men. The presence

²¹See columns (1) and (2) Tables 13 and 14 in the Appendix.

of any sisters has a significant and positive impact on women's education, however, the presence of any brothers has a significant and negative effect on women's education.²² Also the significance of the variable 'any brothers' is very weak and fragile. Indeed, if one introduces a variable such as 'number of siblings squared' the variable 'any brothers' is not significant any more. The effect of the presence of any sisters or any brothers on men's education is not significant. This conclusion contradicts the results observed by Bauer and Gang (2001), who find that the presence of only sisters has a significant negative effect on West German males.

Table 5: Effects of the sibling sex composition on educational attainment

Column	(3)	(4)	(5)
Variable	Coef.	Coef.	Coef.
Age	0.005	0.005	0.005
Father's Education	0.200***	0.200***	0.200***
Mother's Education	0.196***	0.196***	0.197***
Sex	-0.326***	-0.328***	-0.328***
Number siblings	-0.183***	-0.149***	-0.162***
Any sisters	0.248***		
Any brothers		-0.159*	
Percent of sisters			0.003***
Constant	8.545***	8.728***	8.507***
R^2	0.235	0.234	0.235
N	2935	2935	2935

The occupational status of the father is also used in the regressions.

Significance Level: * = $p < 10\%$, ** = $p < 5\%$, *** = $p < 1\%$.

4.2.2 The Matching Results

In order to control for unobserved heterogeneity, two types of matching procedures are applied, the one-to-one matching with the nearest neighbor matching and the kernel based matching using the gaussian kernel and the Epanechnikov kernel, respectively. First a probit model is estimated to obtain the propensity score. Then, individuals who have at least a sister (or a brother) are matched with somebody with as close characteristics as possible, but who does not have any sisters (or any brothers). The differences in educa-

²²See columns (3) to (5) Tables 13 and 14 in the Appendix.

tional attainment give us the estimated effects of having at least a sister or a brother. The estimates, obtained for the treatment on the treated in Table 6 (or Table 7), provide the estimated effects of having at least a sister (or a brother) for those who actually have a sister, and the estimates, obtained for the treatment on the untreated in the Table 6 (or Table 7), show the estimated effects of having at least a sister (or a brother) for those who have not any. To check the significance of the results, the bootstrapping method is used to obtain estimates of the standard deviations.

Table 6: Effects of sisters on educational attainment with matching

The matching estimators	Treated	Controls	Difference	St.E*
Average Treatment on the Treated				
Nearest-neighbor	11.990	11.718	0.272	
Gaussian kernel	11.990	11.633	0.357	(0.090)
Epanechnikov kernel	11.990	11.596	0.394	(0.098)
N		2131		
Average Treatment on the Untreated				
Nearest-neighbor	12.103	12.196	0.093	
Gaussian kernel	12.103	12.270	0.166	(0.133)
Epanechnikov kernel	12.103	12.230	0.127	(0.133)
N		1719		

*Bootstrapping: 300 replications.

Table 7: Effects of brothers on educational attainment with matching

The matching estimators	Treated	Controls	Difference	St.E*
Average Treatment on the Treated				
Nearest-neighbor	11.917	11.857	0.061	
Gaussian kernel	11.917	11.860	0.057	(0.097)
Epanechnikov kernel	11.917	11.801	0.116	(0.100)
N		2212		
Average Treatment on the Untreated				
Nearest-neighbor	12.207	12.186	-0.021	
Gaussian kernel	12.207	12.197	-0.011	(0.124)
Epanechnikov kernel	12.207	12.178	-0.029	(0.137)
N		1638		

*Bootstrapping: 300 replications.

We see that the estimated return to having at least a sister (or a brother) for those who actually had at least a sister (or a brother) is around 0.27 to 0.39 years of education (or 0.06 to 0.12). The estimated return to having at least a sister (or a brother) for those who had not a sister (or a brother) is around 0.09 to 0.16 years of education (or -0.03 to -0.01). The results show that the effects to have a sister for those who had at least a sister are significant, but the effects to have a sister or a brother for individuals who do not have any are not significant. Moreover, the results suggest that if individuals who do not have at least one brother had one of them, the effects on their educational attainment would be weaker, and even negative, but these results are not significant.

5 Effect of an elder sibling

The last step of the empirical analysis is based on a sample which is extracted from the full sample used in the previous studies. This new sample is only composed of couples of siblings. To identify siblings, the identification number of the mother is used. Individuals with the same mother are considered as siblings. This sibling sample consists of 292 observations, i.e. 146 pairs. One difference between the two samples is the age of the sample persons. The sibling sample is younger because this sample is composed of households from the socio-economic panel. As the identification number of the mother is used to identify siblings, it is necessary that the mother was also interviewed by the socio-economic panel. Average ages are 29.09 for the siblings sample and 37.85 for the full sample. Due to the fact that this sample is younger, the average level of education is higher. Educational attainment of both parents are also higher for this sibling sample. Summary statistics are described in Table 8.

Table 8: Summary statistics: Sibling Sample

Variable	Mean	St.D.
The Sibling Sample		
Educational attainment (in Years)	12.315	2.540
Education	2.671	1.002
Age	29.092	4.576
Educational attainment: mother	10.364	1.644
Educational attainment: father	11.240	2.264
Missing	0.079	0.270
Farmer	0.027	0.163
Worker	0.336	0.473
Employee	0.058	0.234
Middle	0.277	0.448
Superior	0.140	0.348
Self-employed	0.082	0.275
Sex	0.377	0.485
Elder schooling	12.483	2.621
Elder age	30.808	4.563
Elder sex	0.336	0.474
Younger schooling	12.147	2.454
Younger age	27.377	3.908
Younger sex	0.418	0.495

5.1 Results

To run the ordered probit, only the youngest sample is considered, therefore there are only 146 observations. Information on the eldest child are used as explanatory variables.

Table 9: Determinants of final educational attainment

Variable	Coef.	St.E.
Elder years of schooling	0.287***	(0.050)
Elder sex	-0.095	(0.215)
Elder age	0.085*	(0.043)
Educational attainment: father	0.288***	(0.068)
Educational attainment: mother	0.066	(0.087)
Sex	-0.012	(0.210)
Age	-0.001	(0.050)
Number of siblings	-0.164	(0.120)
Cut1	7.335	(1.652)
Cut2	9.631	(1.687)
Cut3	11.369	(1.765)
Cut4	11.910	(1.791)
Log-likelihood		-128.245
	$\chi^2(13)$	$p > \chi^2$
LR test	97.902	0.000
Number of Observations N = 146		
Significance Level:*= $p < 10\%$, **= $p < 5\%$, ***= $p < 1\%$.		

The results of the ordered probit model are depicted in Table 9. The results are not sufficiently robust, because there are not enough observations in the sample. However, the results show that the educational attainment of an elder child has a significant and positive effect on educational decisions of the second child. Moreover, one can notice that the educational attainment of the mother, compared to the full sample results, is not significant any more.

6 Conclusion

In this paper, using the GSOEP data, the relation between sibling composition and educational attainment in West Germany is analyzed. The following conclusions have been reached.

The number of children in the family has a negative effect on educational attainment for families with at least two children. Children in large families perform worse than children with fewer siblings, all else equal.

The number of brothers in the family has a negative effect on educational attainment. A possible interpretation is that West German parents tend to favor the education of the boys. They invest less in girls. The sibling sex composition has a significant effect on the educational attainment's of women. To have at least one sister plays an important role on the educational attainment of women, however, the sibling composition has no effect on men's educational attainment.

The analysis of families backgrounds' characteristics assumes unobservable heterogeneity. In order to control for this heterogeneity, the matching method is used. The results suggest that the effect of having at least one sister (or one brother) is weaker (or negative) for the individuals from the group who do not have one sister (or one brother) compared with individuals from the group who have already at least one sister (or one brother).

Finally, using a sibling sample, the effect of the educational attainment of an elder sibling on the educational decisions of a younger child is accounted. The results show a significant and positive effect.

The paper shows that the sibling characteristics (composition and educational attainment) have a decisive effect on educational attainment of individuals, but these effects and their importance differ across the sex of the respondents.

Further steps of the analysis will consist in using longitudinal data, the introduction of a multiple treatment matching method and, finally, the analysis of the implication of the marital status of parents on the educational attainment of children.

7 Appendix

Table 10: Variables used in the analyses.

Variable	Description	Range
School	educational attainment (in years)	7-18
Sex	woman	0-1
Age	1996 - year of birth	22-56
Cohort1	year of birth<1951	0-1
Cohort2	year of birth \in (1951;1961)	0-1
Cohort3	year of birth>1961	0-1
Education	corresponding to the education level	1-5
The sibling composition:		
Any sisters	at least one sister	0-1
Any brothers	at least one brother	0-1
Number of sisters	number of sisters	0-11
Number of brothers	number of brothers	0-11
Number of siblings	number of siblings	0-14
Percentage of sisters	percentage of sisters in the sibship*	0-1
Educational attainment: the respondent has...		
No degree	no degree	0-1
Low school degree	low school degree	0-1
Apprenticeship	an apprenticeship certificate	0-1
Technical College	technical degree	0-1
Interschool	intermediate degree	0-1
Professional college	advanced technical college certificate	0-1
High school	high school degree	0-1
Higher Technical College	high technical degree	0-1
University	university degree	0-1

*Including the respondent.

to be continued...

Table 10 continued...

Variable	Description	Range
The Education of the Mother		
No info	no information	0-1
University Degree	university degree	0-1
Educational attainment**	educational attainment (in years)	7-18
The Education of the Father		
No info	no information	0-1
University Degree	university degree	0-1
Educational attainment**	educational attainment (in years)	7-18
The Occupation Status of the father: the father is...***		
Missing	no info	0-1
Farmer	a farmer	0-1
Worker	a worker	0-1
Employee	an employee	0-1
Middle Position	a middle position	0-1
Senior Position	a senior position	0-1
Self-employed	self employed	0-1
The Sibling Sample		
Elder schooling	educational attainment of the oldest child	7-18
Elder age	age of the oldest child	23-47
Elder sex	sex of the oldest child	0-1
Age	age of the second child	22-46
Sex	sex of the second child	0-1

** Own calculations.

***These categories are presented by Lauer (2002).

Table 11: Relationship between sibship and educational attainment

The respondent is a man				
Number of siblings	Gender sib.	Mean	(St.D.)	Obs.
0 sibling		12.23	(2.59)	512
1 sibling		12.60	(2.62)	690
	0 sist.	12.66	(2.67)	354
	1 sist.	12.53	(2.58)	336
2 siblings		12.23	(2.66)	506
	0 sist.	11.74	(2.43)	120
	1 sist.	12.32	(2.63)	288
	2 sist.	12.61	(2.94)	98
3 siblings		11.90	(2.41)	217
	0 sist.	10.98	(1.87)	27
	1 sist.	12.17	(2.57)	73
	2 sist.	12.13	(2.64)	83
	3 sist.	11.53	(1.56)	34

Table 12: Relationship between sibship and educational attainment

The respondent is a woman				
Number of siblings	Gender sib.	Mean	(St.D.)	Obs.
0 sibling		11.99	(2.36)	403
1 sibling		12.21	(2.47)	516
	0 sist.	12.03	(2.31)	270
	1 sist.	12.41	(2.63)	246
2 siblings		11.81	(2.43)	376
	0 sist.	11.63	(2.44)	93
	1 sist.	11.95	(2.52)	203
	2 sist.	11.65	(2.17)	80
3 siblings		11.40	(2.11)	169
	0 sist.	11.35	(1.77)	24
	1 sist.	11.40	(2.14)	63
	2 sist.	11.61	(2.35)	68
	3 sist.	10.50	(0.98)	14

Table 13: Effects of the sibship sex composition on men educational attainment

Column Variable	(1) Coef.	(2) Coef.	(3) Coef.	(4) Coef.	(5) Coef.
Age	0.027***	0.027***	0.027***	0.028***	0.027***
Father's Education	0.225***	0.225***	0.225***	0.225***	0.225***
Mother's Education	0.181***	0.181***	0.181***	0.181***	0.181***
Number siblings	-0.172***	-0.244***	-0.185***	-0.166***	-0.172***
Number sisters		0.144**			
Any sisters			0.151		
Any brothers				-0.076	
Percent of sisters					0.002
Constant	6.638***	6.641***	7.273***	7.374***	7.252***
R^2	0.242	0.244	0.242	0.242	0.242
N	1678	1678	1678	1678	1678

The occupational status of the father is also used in the regressions.

Significance Level:*= $p < 10\%$, **= $p < 5\%$, ***= $p < 1\%$.

Table 14: Effects of the sibship sex composition on women educational attainment

Column Variable	(1) Coef.	(2) Coef.	(3) Coef.	(4) Coef.	(5) Coef.
Age	-0.024***	-0.025***	-0.025***	-0.023***	-0.024***
Father's Education	0.148***	0.154***	0.150***	0.150***	0.152***
Mother's Education	0.225***	0.225***	0.228***	0.229***	0.229***
Number siblings	-0.131***	-0.237***	-0.165***	-0.109***	-0.133***
Number sisters		0.206***			
Any sisters			0.381***		
Any brothers				-0.250*	
Percent of sisters					0.004***
Constant	8.562***	8.526***	9.089***	9.381***	9.035***
R^2	0.236	0.240	0.241	0.238	0.241
N	1257	1257	1257	1257	1257

The occupational status of the father is also used in the regressions.

Significance Level:*= $p < 10\%$, **= $p < 5\%$, ***= $p < 1\%$.

References

- Ashenfelter, O. and Zimmerman, D. J.: 1997, Estimates of the returns to schooling from sibling data: fathers, sons and brothers, *The Review of Economics and Statistics* **79**(1), 1–9.
- Bauer, T. K. and Gang, I. N.: 2001, Sibling rivalry in educational attainment: The German case, *Labour* **15**(2), 237–255.
- Becker, G. S.: 1964, *Human Capital*, The University of Chicago Press.
- Becker, G. S.: 1981, *A Treatise on the Family*, Harvard University Press.
- Becker, G. S. and Tomes, N.: 1976, Child endowments and the quantity and quality of children, *The Journal of Political Economy* **84**(4), S143–S162.
- Bjoerklund, A. and Jaentti, M.: 1998, The impact of the number of siblings on mens adult earnings: evidence from Finland, Sweden and the United States, *Working Paper* .
- Blundell, R., Dearden, L. and Sianesi, B.: 2001, Estimating the returns to education: Models, methods and results, *Paper presented at the RSS Social Statistics Section* .
- Blundell, R., Dearden, L. and Sianesi, B.: 2003, Evaluating the impacts to education on earnings in the UK: Models, methods and results from the NCDS, *The Institute for Fiscal Studies, Working paper* (03/20).
- Brody, G. H.: 1996, *Sibling relationships: Their causes and consequences*, Ablex Publishing Corporation.
- Butcher, K. F. and Case, A.: 1994, The effect of sibling sex composition on women’s education and earnings, *The Quarterly Journal of Economics* **109**(3), 531–563.
- Conley, D.: 2000, Sibship sex composition: Effects on educational attainment, *Social Science Research* **29**, 441–457.
- Goedde, I. and Schnabel, R.: 1998, Does family background matter? - Returns to education and family characteristics in Germany, *Discussion Paper* (98-60).

- Grawe, N. D. and Mulligan, C. B.: 2002, Economic interpretations of inter-generational correlations, *Journal of Economic Perspectives* **16**(3), 45–58.
- Greene, W. H.: 2003, *Econometric Analysis*, fifth edn, Pearson Education Inc.
- Hauser, R. M. and Kuo, H.-H. D.: 1995, Does the gender composition of sibship affect educational attainment, *CDE Working Paper* (95-06).
- Heckman, J. J., Ichimura, H. and Todd, P. E.: 1998, Matching as an econometric evaluation estimator, *The Review of Economic Studies* **65**(2), 261–294.
- Heer, D. M.: 1985, Effects of siblings number on child outcome, *Annual Review of Sociology* **11**, 27–47.
- Iacovou, M.: 2001, Family composition and children’s educational outcomes, *Working Paper* .
- Kaestner, R.: 1996, Are brothers really better? Sibling sex composition and educational achievement revisited, *NBER Working Paper* **5521**.
- Lauer, C.: 2002, Family background, cohort and education: A French-German comparison, *ZEW Discussion Paper* (02-12).
- Lechner, M.: 2002a, Programme heterogeneity and propensity score matching: An application to the evaluation of active labor market policies, *The Review of Economics and Statistics* **84**(2), 205–220.
- Lechner, M.: 2002b, Some practical issues in the evaluation of heterogeneous labor market programmes by matching methods, *Journal of the Royal Statistical Society, Serie A* **165**(1), 59–82.
- Lindahl, L.: 2002, Do birth order and family size matter for intergenerational income mobility? Evidence from Sweden, *Working Paper from Swedish Institute for Social Research* **5/2002**.
- Maddala, G.: 1999, *Limited-dependent and Qualitative Variables in Econometrics*, Cambridge University Press.

- Olneck, M. R. and Bills, D. B.: 1979, Family configuration and achievement: Effects of birth order and family size in a sample of brothers, *Social Psychology Quarterly* **42**(2), 135–148.
- Plug, E.: 1999, Is schooling a family thing? Effects of grandparents, parents, brothers and sisters on the school choices of boys and girls, *Discussion Paper*.
- Powell, B. and Steelman, L. C.: 1989, The liability of having brothers: paying for college and the sex composition of the family, *Sociology of Education* **62**(2), 134–147.
- Powell, B. and Steelman, L. C.: 1990, Beyond sibship size: Sibling density, sex composition and educational outcomes, *Social Forces* **69**(1), 181–206.
- Rosenbaum, P. R. and Rubin, D. B.: 1983, The central role of the propensity score in observational studies for causal effects., *Biometrika* **70**(1), 41–55.
- Solon, G., Corcoran, M., Gordon, R. and Laren, D.: 1987, The effect of family background on economic status: A longitudinal analysis of sibling correlations, *Working Paper* (2282).
- Stafford, F. P.: 1987, Women’s work, sibling competition, and children’s school performance, *The American Economic Review* **77**(5), 972–980.
- Wells, T.: 1995, Does family background affect educational attainment differently according to family structure, birth order and sex, *National survey of family and households, Working paper* (70).