



UvA-DARE (Digital Academic Repository)

Family background and children's schooling outcomes

de Haan, M.

[Link to publication](#)

Citation for published version (APA):

de Haan, M. (2008). Family background and children's schooling outcomes Amsterdam: Thela Thesis

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <http://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

FAMILY BACKGROUND AND CHILDREN'S SCHOOLING OUTCOMES

ISBN 978 90 361 0093 9

Cover design: Crasborn Graphic Designer bno, Valkenburg a.d. Geul

This book is no. 441 of the Tinbergen Institute Research Series, established through cooperation between Thela Thesis and the Tinbergen Institute. A list of books which already appeared in the series can be found in the back.

FAMILY BACKGROUND AND CHILDREN'S SCHOOLING OUTCOMES

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
Prof. dr. D.C. van den Boom
ten overstaan van een door het college voor promoties
ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op donderdag 4 december 2008, te 14:00 uur

door

Monique de Haan

geboren te Naarden

PROMOTIECOMMISSIE

Promotor:

Prof. dr. E.J.S. Plug

Overige leden:

Prof. dr. J. Hartog

Prof. dr. H. Maassen van den Brink

Prof. dr. K.G. Salvanes

Dr. E. Leuven

Dr. B. van der Klaauw

Faculteit der Economie en Bedrijfskunde

Acknowledgements

Five years ago, when I was studying economics at the University of Amsterdam, I followed a course called "onderzoekspracticum". It was the first course in which I did not have to study a book or lecture notes, but instead I was supposed to do my own empirical research project. I really enjoyed working with a data set, writing and presenting my results and I decided that I wanted to continue doing empirical research and to write a PhD thesis. During the years I learned though that writing a thesis is a lot more challenging than writing a paper for that course. It is a process which comes with ups and downs, but I always enjoyed doing empirical research and I'm happy that I decided to become a PhD-student. Of course this thesis would not be here without the help and support of colleagues, family and friends and I would like to take this opportunity thank them.

First I would like to thank Erik and Hessel. I want to thank Erik for being my supervisor, for always carefully reading my work, for giving useful suggestions and for letting me use our joint work as part of this thesis. I have learned a lot and I enjoyed working with you. I want to thank Hessel for reading and commenting on my work, for giving advice, and for giving useful suggestions which helped to improve my papers. Although officially I have only one supervisor, I have been fortunate to learn from two supervisors.

I further want to thank all other colleagues at the tenth and eleventh floor for the discussions in the hallways, the coffee breaks, the labour seminars, the reading groups and the lunches. During the years that I've spent on writing my thesis, the composition and number of colleagues have changed quite a bit. Initially I wanted to put the names here of those I want to thank in particular and I started making a list of names. But after some time I gave up.....there are just too many people that helped me at some point during the process of writing my thesis, by giving helpful suggestions, cheering me up, joining me at conferences, asking (difficult) questions at seminars, joining discussions in front of the coffee machine, etc.

There are however two 'groups' I want to thank in particular. First I want to thank everyone who joined the lunches. Sometimes I had lunch with 12 people (all around one small table) and sometimes with two or three (which is actually a better fit for the table). But no matter how big the group was, I always enjoyed the lunches. I want to thank all the co-lunchers for the laughs, the useful (and less useful) discussions and for providing a fun and energy-giving break of the day. Second I would like to thank everyone who participated in the reading groups in which we discussed our own work or other papers related to our research. I always enjoyed these meetings and I learned a lot.

Finally I would like to thank my family and friends. I want to thank my brother Thomas for all the discussions we had about economics, especially when we were both still living with our parents. I want to thank my parents for the times they had to listen to these discussions, but most of all for being there for me and for supporting me at all times. Above all I want to thank Johan for loving and supporting me, for listening to my stories about economics, econometrics and the crazy world of scientists, for encouraging me to finish this thesis and for always being there for me.

Contents

Introduction	1
1 The effect of parents' schooling on child's schooling: A nonparametric bounds analysis	5
1.1 Introduction	5
1.2 Empirical specification	9
1.3 Data	18
1.4 Results	21
1.5 Conclusion	35
Appendix	37
2 Estimating intergenerational schooling mobility on censored samples: Consequences and remedies	39
2.1 Introduction	39
2.2 Mobility models using censored data	42
2.3 Data	46
2.4 Results	48
2.5 Can we treat expectations as realizations?	53
2.6 Conclusion	62
3 Birth order, family size and educational attainment	65
3.1 Introduction	65

3.2	Theoretical and empirical background	68
3.3	Data	74
3.4	Results	75
3.5	Potential mechanisms behind the birth order effects	84
3.6	Conclusion	94
4	Summary and conclusions	97
	Bibliography	101
	Nederlandse samenvatting	109

Introduction

“It is well documented that individuals are very diverse in a large variety of abilities, that these abilities account for a substantial amount of the interpersonal variation in socioeconomic outcomes, and that this diversity is already apparent at an early age. The family plays a powerful role in shaping these abilities, contributing both genetic endowments and pre- and post-natal environments, which interact to determine the abilities, behavior and talents of children” –Cunha, Heckman, Lochner and Masterov (2006)

What determines the success of children? As stated by the quote from Cunha, Heckman, Lochner and Masterov in their handbook chapter, it is well documented that family background is a very important, if not the most important, determinant of socioeconomic outcomes of children. Not only scientists but also policy makers stress the importance of the family. In 2005 OECD Ministers responsible for social policies met in Paris to discuss "Extending opportunities: how active social policies can benefit us all". The Final Communiqué from this meeting stated that "Social and family policies must help give children and young people the best possible start to their lives and help them to develop and achieve through their childhood into adulthood." It further communicated that "Promoting child development requires society and families to invest adequate resources..... Special effort should be targeted on the families that are struggling to give their children the resources, both financial and time, that they need."

In order to use policy measures to help children and reduce inequalities between chil-

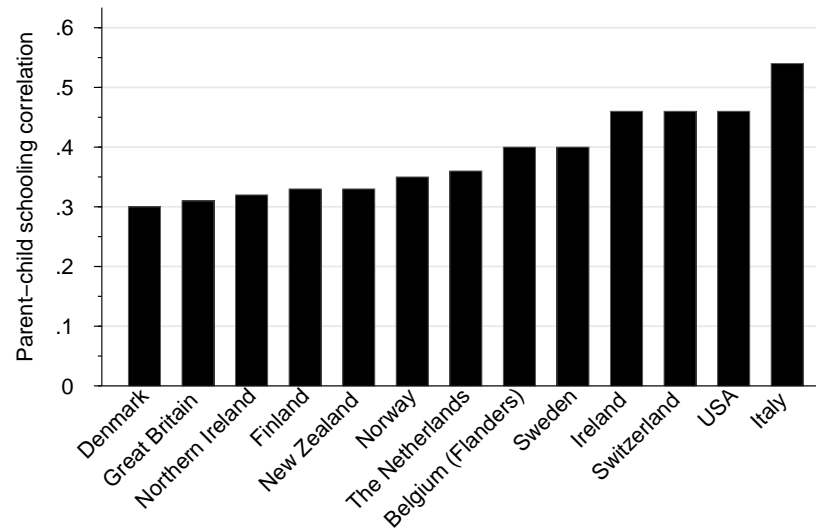
dren from different backgrounds, it is important to know how family background affects the socioeconomic outcomes of children. There are multiple channels through which the family can influence children's outcomes. The different family characteristics are generally strongly related, making it difficult to disentangle the different mechanisms. Disentangling the different mechanisms is not only challenging but also very important. Interpreting an association as a causal effect can lead to wrong conclusions, whereby a particular family characteristic could seem to affect children's socioeconomic outcomes while actually it is picking up the effect of something else.

This thesis will focus on one particular socioeconomic outcome, children's completed years of schooling. It will investigate why some children obtain the highest schooling level while others drop out without a diploma, and focus on what role family characteristics play in these differences in schooling outcomes.

Figure 1 shows the correlation between parents' and child's years of schooling, for 13 different countries, based on results from Hertz *et al.* (2007). This figure shows that children's schooling outcomes are strongly related to the schooling of their parents. These correlation coefficients are however not necessarily prove of a causal effect of parental schooling on child's schooling, it might pick up the effect of related characteristics such as genetic endowments and parent's child rearing talents. The remaining chapters will study the effects of father's and mother's schooling, birth order and family size on children's educational attainment with a particular focus on the use of different techniques to identify the causal impacts of these family background characteristics.

Chapter 1 investigates the effect of parents' schooling on child's schooling. Since observed associations between parental and child's schooling do not necessarily reflect a true causal effect, this chapter applies a relatively new technique, a nonparametric bounds analysis based on a study by Manski and Pepper (2000). The analysis starts with making no assumptions and then adds some relatively weak and testable assumptions to tighten the bounds. The assumptions are relatively weak in the sense that they do not impose a

Figure 1: Parent-child schooling correlations for different countries



source: Based on Hertz *et al.* (2007). The height of the bar presents the long-run average across 5-year cohorts of the correlation coefficient between the average of mother's and fathers schooling and child's schooling. Data set used: International Adult Literacy Survey except for the USA: International Social Survey Programme

linear effect of parents' schooling, they allow for a potential positive correlation between parents' schooling and unobserved endowments and they allow for possible interaction effects between heritable endowments and parents' level of schooling. Although the bounds on the effect of parents' schooling include a zero effect, the upper bounds are informative especially for the effect of increasing parents' schooling from a high school degree to a bachelor's degree. Both for the effect of mother's schooling as for the effect of father's schooling the nonparametric upper bounds are significantly lower than the OLS results.

Chapter 2 also studies the impact of parental schooling on child's schooling. The focus of this chapter is on the problem that many intergenerational mobility studies use samples in which part of the children is still in school. It investigates the consequences of this often encountered censoring problem and evaluates three solutions to it: maximum likelihood approach, replacement of observed with expected years of schooling, and elimination of all school-aged children. This chapter test how the three correction methods deal with

censored observations. The main finding is that the method that treats parental expectations as if they were realizations seems to fix the censoring problem quite well, while the other methods produce a (small) positive bias in the estimates.

Chapter 3 does not study the impact of parental schooling but instead investigates two other family background components, the effect of the number of children and birth order on children's completed years of schooling. An instrumental variables approach is used to identify the effect of family size. Instruments for the number of children are twins at last birth and the sex mix of the first two children. The effect of birth order is identified, by examining the relation with years of education for different family sizes separately. No significant effect of the number of children on educational attainment of the oldest child is found, while birth order has a significant negative effect. This decline in schooling outcomes with birth order turns out to be approximately linear. This chapter also investigates potential mechanisms behind the birth order effects; child-spacing and the allocation of parental resources. The results show that the negative birth order effects do not vary with the average age gap between children. Information on financial transfers to children shows that earlier born children have a higher probability of receiving money from their parents than later born children, also the amount they receive is higher. These results indicate that the allocation of parental resources is a potential mechanism behind the birth order effects.

Finally, Chapter 4 will summarize the main findings and conclusions of the preceding chapters.

Chapter 1

The effect of parents' schooling on child's schooling: A nonparametric bounds analysis

1.1 Introduction

Is there an effect of parents' schooling on the schooling of their child? This question has received much attention in the empirical literature. Most if not all studies find a positive association between parental and child's schooling. Haveman and Wolfe (1995) state in a survey of the literature

"...perhaps the most fundamental economic factor is the human capital of parents, typically measured by the number of years of schooling attained. This variable, emphasized in the earlier studies of the intergenerational transmission of socioeconomic status, is included in virtually every study described {in this review}; it is statistically significant and quantitatively important, no matter how it is defined."– Haveman and Wolfe (1995, pp.1855).

Most of the studies discussed in the overview regress child's schooling on the schooling of the parents. To give a causal interpretation to the results of these regressions, one has to impose a number of assumptions; a linear impact of parental years of schooling and

no correlation between parents' schooling and unobserved endowments affecting their child's schooling. Since these are rather strong assumptions, the positive association does not need to be a true causal relation. In an attempt to isolate the causal impact of parents' schooling, different identification strategies have been applied in the recent empirical literature.

One of the approaches is to use a sample of identical twins to eliminate the correlation between parental schooling and child's schooling attributable to genetics (Behrman and Rosenzweig (2002, 2005), Antonovics and Goldberger (2005)). These studies find a positive and significant relation between both father's and mother's schooling and the schooling of their child. However, the within-twin estimates, whereby they difference out the genetic factors, indicate that the effect of parents' schooling is lower than the OLS estimates and that this decline is strongest for mothers.

A second approach uses a sample of adoptees, whereby they exploit the fact that there is no genetic link between adoptive parents and their adopted child (Björklund *et al.* (2006), Sacerdote (2002, 2007), Plug (2004)). The main findings of these adoption studies are that the estimates of the relation between parents' and child's schooling is significantly smaller when estimated on a sample of adoptees instead of on a sample of own birth children. This indicates that a large part of the intergenerational association is due to genetic transmission of endowments.

A third identification strategy is an instrumental variable (IV) approach. By using a change in compulsory schooling laws in Norway as instrument, Black *et al.* (2005A) find insignificant effects of parental schooling on child's years of schooling, except for the effect of mother's schooling on the schooling of her son. Chevalier (2004) uses a change in the minimum school leaving age in Britain and finds that the effects of parents' schooling on the probability that the child of the same gender has post-compulsory schooling is positive and higher than the results without using an instrument. Oreopoulos *et al.* (2006), Carneiro *et al.* (2007) and Maurin and McNally (2008) focus on the effect of parents'

schooling on intermediate schooling outcomes. They all find a significant impact of parents' schooling and most of the IV estimates are somewhat higher than the OLS results.

These identification methods generally put strong requirements on the data, since one needs a data set that includes information on both parents' and child's completed schooling, that includes a large enough sample of twins or adoptees, or that includes good instruments for schooling, which are scarce. And even when these rich data sources are available one still has to impose a number of assumptions to be able to use these methods to say something about the causal impact of parents' schooling.

The method using within-twin differences to identify the effect of parents' schooling strongly relies on the assumption of a linear impact of parents' years of schooling. This implies that an extra year of primary education should have the same effect as an extra year of university education. But is this plausible? Another assumption that these twin studies have to make is that there are no interaction effects between genetic endowments and the schooling level of the parents. These studies further assume that monozygotic twin mothers and fathers have identical unobserved endowments and that all differences in schooling levels between these twin parents are exogenous. This has been questioned for example by Bound and Solon (1999).

Studies using samples of adoptees assume that adoptees are randomly assigned to their adoptive parents. This assumption might be violated when adoption agencies match adoptees to adoptive parents on the basis of characteristics of the biological parents. Another assumption which is necessary to give a causal interpretation to the results is that parents' child rearing talents must be uncorrelated to their level of schooling. Also many of these studies rule out potential interaction effects between heritable endowments and the environment in which the children are raised, something which has been criticized in the literature (Cunha and Heckman (2007)).¹

¹Björklund *et al.* (2006) use a sample of Swedish adoptees to estimate intergenerational mobility. They have information on the adoptive and biological parent of the adoptees and include an interaction term in their analysis. They find some evidence for a positive interaction effect between pre-birth factors and post-birth environment.

When using an instrumental variable approach one generally does not have to impose these assumptions, but instrumental variables are often only able to identify a local average treatment effect. Also whether one can interpret the results of IV studies as causal depends on the validity of the instruments something which can, unfortunately, not be tested.

This chapter uses a different approach to investigate the effect of parents' schooling on child's schooling; a nonparametric bounds analysis based on Manski and Pepper (2000), using the most recent version of the Wisconsin Longitudinal Study. We start with making no assumptions and then add much weaker and testable assumptions to tighten the bounds. The assumptions are much weaker in the sense that they do not impose a linear effect of parents' schooling, they allow for a potential positive correlation between parents' schooling and unobserved endowments and they allow for possible interaction effects between heritable endowments and parents' level of schooling. Also, in contrast to most instrumental variable studies, we are able to identify bounds on the effect of parents' schooling over the entire schooling distribution.

There is a trade-off between making less strong (and more credible) assumptions and the information one obtains about the effect of interest. This chapter will obtain bounds instead of point identification. The contribution of this chapter is that it makes relatively weak and testable assumptions, while the identification strategies mentioned above are based on much stronger assumptions and give point estimates which are only informative if these assumptions are correct. And, although there have been studies applying a nonparametric bounds analysis (Gerfin and Schellhorn (2006), González (2005), Pepper (2000), Lechner (1999), Blundell *et al.* (2007)), this chapter is the first study investigating intergenerational schooling mobility using a nonparametric bounds analysis.

The remainder of the chapter is organized as follows. Section 1.2 gives the empirical specification. Section 1.3 gives a description of the data. Section 1.4 will give the results of the nonparametric bounds analysis and compares them to the results of an exogenous treatment selection assumption. And finally Section 1.5 will summarize and conclude.

1.2 Empirical specification

For each child we have a response function $y_i(\cdot) : T \rightarrow Y$ which maps treatments $t \in T$ into outcomes $y_i(t) \in Y$. Where the treatment t is the level of schooling of the parent and y is years of schooling of the child. For each child we observe the realized level of parental schooling z_i and his realized years of schooling $y_i \equiv y_i(z_i)$, but we do not observe the potential outcomes $y_i(t)$ for $t \neq z_i$. To simplify notation the subscript i will be dropped in the following.

We are interested in the mean effect of an increase in parental schooling from s to t on child's schooling, that is

$$\Delta(s, t) = E[y(t)] - E[y(s)] \quad (1.1)$$

By using the law of iterated expectations and the fact that $E[y(t)|z = t] = E[y|z = t]$ we can write

$$E[y(t)] = E[y|z = t] \cdot P(z = t) + E[y(t)|z \neq t] \cdot P(z \neq t) \quad (1.2)$$

With a data set where we observe the schooling of a child and his parent we can observe the mean schooling of a child whose parent has schooling level t and the probability that the parent has schooling level t . However, for a child with a parent who does not have schooling level t we cannot observe what his mean schooling would have been if his parent would have had schooling level t . That is, we cannot observe $E[y(t)|z \neq t]$. It is only possible to say more about the effect of interest by augmenting the things that are observed with assumptions.

Manski (1989) shows though that it is possible to identify bounds on $E[y(t)]$ without making any assumptions if the support of the dependent variable is bounded, which is the case with child's schooling. By substituting $E[y(t)|z \neq t]$ by the lowest possible level of education \underline{y} we obtain a lower bound on $E[y(t)]$ and by replacing it with the highest possible level of schooling \bar{y} we obtain the upper bound. This gives Manski's no-assumption

bounds (1989)

No-assumption bounds

$$\begin{aligned} E[y|z = t] \cdot P(z = t) + \underline{y} \cdot P(z \neq t) \\ \leq E[y(t)] \leq \end{aligned} \tag{1.3}$$

$$E[y|z = t] \cdot P(y = t) + \bar{y} \cdot P(z \neq t)$$

To tighten these no-assumption bounds we will subsequently add the monotone treatment response assumption (MTR) and the monotone treatment selection assumption (MTS) which are introduced and derived in Manski (1997) and Manski and Pepper (2000).

The monotone treatment response assumption states that a child's schooling is weakly increasing in conjectured schooling of his parent:

$$t_2 \geq t_1 \Rightarrow y(t_2) \geq y(t_1) \tag{1.4}$$

This assumes that having a higher educated parent never decreases a child's schooling, which is also suggested by human capital theory (Becker and Tomes (1979), Solon (1999)). A zero effect is not ruled out by this assumption. The MTR assumption implies the following

$$\text{for } u < t \quad E[y(t)|z = u] \geq E[y(u)|z = u] = E[y|z = u]$$

$$\text{so } E[y(t)|z = u] \in [E[y|z = u], \bar{y}]$$

$$\text{for } u > t \quad E[y(t)|z = u] \leq E[y(u)|z = u] = E[y|z = u]$$

$$\text{so } E[y(t)|z = u] \in [\underline{y}, E[y|z = u]]$$

A sample of children and their parents can be divided into three groups; (1) children with a parent that has a schooling level lower than t , (2) those that have a parent with a schooling level equal to t , and (3) children who have a parent with a schooling level higher than t . For the second group we observe the effect on mean schooling of having a parent with schooling level t . For the first group we know that under the MTR assumption their observed mean schooling is less than or equal to what their mean schooling would have been if their parent did have schooling level t . So we can use the mean schooling we observe for this first group to tighten the lower bound. For the third group we know that if they would have had a parent with schooling level t , their mean schooling would have been lower than or equal to their current mean schooling. We can therefore use the mean schooling we observe for this third group to tighten the upper bound. By combining this with the no-assumption bounds above we get the MTR bounds:

MTR bounds

$$\begin{aligned}
 E[y|z < t] \cdot P(z < t) + E[y|z = t] \cdot P(z = t) + \underline{y} \cdot P(z > t) \\
 \leq E[y(t)] \leq
 \end{aligned}
 \tag{1.5}$$

$$\bar{y} \cdot P(z < t) + E[y|z = t] \cdot P(z = t) + E[y|z > t] \cdot P(z > t)$$

To narrow the bounds we will add the monotone treatment selection assumption. Under this assumption children with higher schooled parents have weakly higher mean schooling functions than those with lower schooled parents:

$$u_2 \geq u_1 \Rightarrow E[y(t)|z = u_2] \geq E[y(t)|z = u_1]
 \tag{1.6}$$

This assumption is consistent with higher schooled parents having higher heritable and

child-rearing endowments which can positively (but not negatively) affect their child's schooling. Under the combined MTR-MTS assumption the following holds.²

$$\text{for } u < t \quad E[y(t)|z = t] \geq E[y(t)|z = u] \geq E[y(u)|z = u]$$

$$\text{so } E[y(t)|z = u] \in [E[y|z = u], E[y|z = t]]$$

$$\text{for } u > t \quad E[y(t)|z = t] \leq E[y(t)|z = u] \leq E[y(u)|z = u]$$

$$\text{so } E[y(t)|z = u] \in [E[y|z = t], E[y|z = u]]$$

We can again divide the sample into three groups, children who have a parent with a schooling level lower than t , equal to t , or higher than t . If the schooling of the parents of the first group would be increased to t , we know by the MTS assumption that the mean schooling of the children would be weakly lower than the mean schooling we observe for the children who currently have a parent with schooling level t . We can therefore use the mean schooling we observe for the children who have a parent with schooling level t as an upper bound on the treatment effect for the first group. Similarly we can use it as a lower bound on the treatment effect for the third group. By combining the monotone treatment response assumption and the monotone treatment selection assumption we get the MTR-MTS bounds:³

²The first inequalities follow from the MTS assumption and the second inequalities from the MTR assumption.

³For a full derivation of the MTR and MTR-MTS bounds see Manski (1997) and Manski and Pepper (2000).

MTR-MTS bounds

$$\begin{aligned}
& E[y|z < t] \cdot P(z < t) + E[y|z = t] \cdot P(z = t) + E[y|z > t] \cdot P(z > t) \\
& \leq E[y(t)] \leq \\
& E[y|z = t] \cdot P(z < t) + E[y|z = t] \cdot P(z = t) + E[y|z > t] \cdot P(z > t)
\end{aligned} \tag{1.7}$$

It is possible to test the combined MTR-MTS assumption. Under the MTR-MTS assumption the following holds

$$\text{for } u_2 > u_1$$

$$E[y|z = u_2] = E[y(u_2)|z = u_2] \geq E[y(u_2)|z = u_1] \geq E[y(u_1)|z = u_1] = E[y|z = u_1]$$

So under the MTR-MTS assumption the mean schooling of a child should be weakly increasing in the realized level of schooling of the parent, if this is not the case the MTR-MTS assumption should be rejected.

So far we have obtained bounds on $E[y(t)]$ but we are interested in the effect of an increase in parental schooling ($E[y(t)] - E[y(s)]$). To obtain bounds on this treatment effect we will subtract the lower (upper) bound on $E[y(s)]$ from the upper (lower) bound on $E[y(t)]$ to get the upper (lower) bound. For the bounds using the MTR assumption the lower bound on the effect of an increase in parents' education cannot be negative and is therefore set to zero.

Monotone instrumental variable assumption

Suppose we observe not only the schooling of the child and his parent but also a variable z^* . We could then divide the sample into sub-samples, one for each value of z^* , and for each sub-sample obtain the no-assumption bounds on the basis of equation (1.3). It may well be that the no-assumption bounds are relatively tight for some sub-samples but relatively wide for other sub-samples. We could exploit this variation in the bounds over the sub-samples if z^* satisfies the instrumental variable assumption (Manski and Pepper (2000)). A variable z^* satisfies the instrumental variable assumption, in the sense of mean-independence, if it holds that for all treatments $t \in T$ and all values of the instrument $m \in M$

$$E[y(t)|z^* = m] = E[y(t)] \quad (1.8)$$

This means that the schooling function of the child should be mean-independent of the variable z^* . If z^* satisfies the instrumental variable assumption, we can obtain an IV-lower bound on $E[y(t)]$ by taking the maximum lower bound over all sub-samples and an IV-upper bound by taking the minimum upper bound over all sub-samples. Combining the instrumental variable assumption with the no-assumption bounds gives thus the following IV-bounds

IV-bounds

$$\begin{aligned} & \max_{m \in M} (E[y|z = t, z^* = m] \cdot P(z = t|z^* = m) + \underline{y} \cdot P(z \neq t|z^* = m)) \\ & \leq E[y(t)] \leq \end{aligned} \quad (1.9)$$

$$\min_{m \in M} (E[y|z = t, z^* = m] \cdot P(z = t|z^* = m) + \bar{y} \cdot P(z \neq t|z^* = m))$$

The width of the no-assumption bounds depends on the proportion of children who actually have a parent with schooling level t . The higher $P(z = t)$ the tighter the no-assumption bounds. If for some sub-samples (defined by the values of z^*) the proportion of children who have a parent with schooling level t is higher than for other sub-samples, the no-assumption bounds will be tighter for these sub-samples. The IV-bounds will therefore be tighter than the no-assumption bounds if there is variation in $P(z = t|z^* = m)$ over z^* . This means that the probability that the parent has a certain level of schooling should vary with the value of the instrumental variable.

Since it is difficult to find a variable which satisfies the instrumental variable assumption in equation (1.8) we will use a weaker version; the monotone instrumental variable assumption. A variable z^* is a monotone instrumental variable (MIV) in the sense of mean-monotonicity if it holds that

$$m_1 \leq m \leq m_2 \Rightarrow E[y(t)|z^* = m_1] \leq E[y(t)|z^* = m] \leq E[y(t)|z^* = m_2] \quad (1.10)$$

So instead of assuming mean-independence, the monotone instrumental variable assumption allows for a weakly monotone relation between the variable z^* and the mean schooling function of the child (Manski and Pepper (2000)).

We can again divide the sample into sub-samples on the basis of z^* and obtain no-assumption bounds for each sub-sample. From equation (1.10) it follows that $E[y(t)|z^* = m]$ is no lower than the no-assumption lower bound on $E[y(t)|z^* = m_1]$ and it is no higher than the no-assumption upper bound on $E[y(t)|z^* = m_2]$. For the sub-sample where z^* has the value m we can thus obtain a new lower bound, which is the largest lower bound over all the sub-samples where z^* is lower than or equal to m . Similarly we can obtain a new upper bound by taking the smallest upper bound over all sub-samples with a value of z^* higher than or equal to m .

By repeating this for all $m \in M$ and taking the average we get the following MIV-bounds.

MIV-bounds

$$\begin{aligned} \sum_{m \in M} P(z^* = m) \cdot \left[\max_{m_1 \leq m} \left(\begin{array}{l} E[y|z = t, z^* = m_1] \cdot P(z = t|z^* = m_1) \\ + \underline{y} \cdot P(z \neq t|z^* = m_1) \end{array} \right) \right] \\ \leq E[y(t)] \leq \end{aligned} \tag{1.11}$$

$$\sum_{m \in M} P(z^* = m) \cdot \left[\min_{m_2 \geq m} \left(\begin{array}{l} E[y|z = t, z^* = m_2] \cdot P(z = t|z^* = m_2) \\ + \bar{y} \cdot P(z \neq t|z^* = m_2) \end{array} \right) \right]$$

These MIV bounds are generally wider than the IV bounds. However when the no-assumption upper and lower bounds weakly decrease with the value of z^* , the identifying power of the MIV assumption is as strong as the identifying power of the IV assumption.⁴

Instead of combining the MIV assumption with the no-assumption bounds we can also combine the MIV assumption with the MTR-MTS bounds. This means that instead of obtaining no-assumption bounds for each sub-sample we obtain MTR-MTS bounds for each sub-sample. By replacing the no-assumption bounds in equation (1.11) by the MTR-MTS bounds we obtain the MTR-MTS-MIV bounds.

⁴The no-assumption lower bound is high when $P(z = t)$ is high. So for the lower bound to weakly decrease with the value of z^* , $P(z = t|z^* = m)$ should *decrease* for at least some values of z^* . The no-assumption upper bound is low when $P(z = t)$ is high. So for the upper bound to be weakly decreasing with z^* , $P(z = t|z^* = m)$ should *increase* for at least some values of z^* .

MTR-MTS-MIV bounds

$$\sum_{m \in M} P(z^* = m) \cdot \left[\max_{m_1 \leq m} \begin{pmatrix} E[y|z < t, z^* = m_1] \cdot P(z < t | z^* = m_1) + \\ E[y|z = t, z^* = m_1] \cdot P(z = t | z^* = m_1) + \\ E[y|z > t, z^* = m_1] \cdot P(z > t | z^* = m_1) \end{pmatrix} \right] \leq E[y(t)] \leq$$
(1.12)

$$\sum_{m \in M} P(z^* = m) \cdot \left[\min_{m_2 \geq m} \begin{pmatrix} E[y|z = t, z^* = m_2] \cdot P(z < t | z^* = m_2) + \\ E[y|z = t, z^* = m_2] \cdot P(z = t | z^* = m_2) + \\ E[y|z > t, z^* = m_2] \cdot P(z > t | z^* = m_2) \end{pmatrix} \right]$$

We will use two monotone instrumental variables. The first is the schooling of the grandparent. Since it is unlikely that the schooling function of the child is mean-independent of the schooling of his grandparent we will not use grandparent's schooling as an instrumental variable, but we will use it as a monotone instrumental variable. By using grandparent's schooling as a MIV we assume that the mean schooling function of the child is monotonically increasing (or non-decreasing) in the schooling of the grandparent.

The second MIV is the schooling of the spouse. When we obtain bounds on the effect of mother's schooling we will use the level of schooling of the father as MIV, and if we obtain bounds on the effect of father's schooling we will use the schooling level of the mother as MIV. As with grandparent's schooling the schooling of the spouse is unlikely to satisfy the mean-independence assumption in equation (1.8), we will therefore use it as a MIV and assume that the mean schooling function of the child is non-decreasing in the schooling of the spouse.

Obtaining bounds on the effect of increasing father's/mother's schooling from s to t

works in the same way as was described at the end of the previous subsection. We first obtain the MTR-MTS-MIV upper and lower bounds on $E[y(t)]$ and $E[y(s)]$, and then take the difference between the upper bound on $E[y(t)]$ and the lower bound on $E[y(s)]$ to get the upper bound on $\Delta(s,t) = (E[y(t)] - E[y(s)])$. The lower bound on $\Delta(s,t)$ is set to zero by the monotone treatment response assumption.⁵

1.3 Data

The analysis in this chapter uses the most recent version of the Wisconsin Longitudinal Study (WLS). The WLS is a long-term study based on a random sample of 10,317 men and women who graduated from Wisconsin high schools in 1957. Next to information about the graduates the sample contains comparable data for a randomly selected sibling of most of the respondents. Survey data were collected from the original respondents in 1957, 1964, 1975, 1992, and 2004 and from the selected siblings in 1977, 1994, and 2005. We will mainly use the data from the last two waves (2004, 2005) since these contain updated information about completed schooling of the graduates and their spouses, the selected siblings and their spouses and about the children of both the graduates and the selected siblings. In the last two waves information is collected from 7,265 graduates and 4,271 selected siblings.

The sample that is used in this chapter includes graduates and selected siblings who were married at least once and who have at least one child. It is not possible to link children to spouses, but only possible to link children to respondents and spouses to respondents.

⁵This means that we do *not* use the following assumption to obtain bounds on $\Delta(s,t)$, which is stronger than the assumption in equation (1.10):

$$m_1 \leq m \leq m_2 \Rightarrow E[\Delta(s,t)|z^* = m_1] \leq E[\Delta(s,t)|z^* = m] \leq E[\Delta(s,t)|z^* = m_2] \quad (1.13)$$

Using assumption (1.13) instead of assumption (1.10) would mean that we obtain bounds on the effect of an increase in father's/mother's schooling ($\Delta(s,t)$) for each sub-sample and thus conditional on the monotone instrumental variable. This could be problematic when using the schooling of the spouse as MIV since part of the effect of increasing mother's (father's) schooling could be through the effect that she(he) marries a higher schooled spouse. However, since we do not use assumption (1.13) but instead use assumption (1.10), this is not an issue in the analysis in this chapter.

The sample is therefore further restricted to respondents who only have children from their first marriage to be quite sure that both the spouse and the respondent are the child's biological parents. This gives a final sample of 21,545 children of 5,167 graduates and 2,524 selected siblings.⁶

Information about completed schooling is available in years. For the analysis in this chapter it is necessary to have enough observations for each observed level of parental schooling, therefore we construct schooling variables in levels for the respondents and their spouses as follows:

1	Less than high school	< 12 years
2	High school	12 years
3	Some college	13-15 years
4	Bachelor's degree	16 years
5	Master's degree	17 years
6	More than Master's degree	>17 years

For the schooling of the grandparent we will use the schooling of the head of the household when the parent was 16 (in 80-90% of the cases the father is the head of the household).⁷ We construct schooling variables in levels as with parents' schooling. However since the average schooling level has increased over time we construct different schooling levels for grandparents. The education variable for the head of the household has peaks at

⁶There are some children below the age of 23 and these children might still be in school. In the analysis we eliminate these observations. Chapter 2 shows that when 23% of the sample is censored, eliminating children who are still in school can cause a small positive bias. In the sample in this chapter only 1.5% is below the age of 23. It is unlikely that eliminating these observations can cause a significant bias in the estimates.

⁷Unfortunately this variable is not available for the spouse of the selected sibling

6, 8, 12 and 16 years of schooling, we therefore construct the following schooling variable in levels for the grandparent:

1	Elementary school	≤ 6 years
2	Middle school	7-8 years
3	Some high school	9-11 years
4	Graduated from high school	12 years
5	Some college	13-15 years
6	Bachelor's degree or more	≥ 16 years

Table 1.1 gives some descriptive statistics. Table 1.2 gives mean schooling of the child for each level of mother's and father's schooling and the percentage of observations in each category, and the appendix gives an overview of the educational system in the United States. For more detailed information on the WLS see Sewell *et al.* (2004) and WLS (2006) and the references therein.

Table 1.1: Summary Statistics

	Mean	Std. dev.	N
Years of schooling child	14.50	2.32	21,545
Child has bachelor's degree	0.46	0.50	21,545
Gender (female=1)	0.49	0.50	21,545
Age child	38.34	5.50	21,494
Years of schooling father	13.52	2.70	21,545
Years of schooling mother	13.03	1.86	21,545
Level of schooling father	2.87	1.43	21,545
Level of schooling mother	2.56	1.02	21,545
Schooling head of household when father was 16	9.86	3.40	14,614
Schooling head of household when mother was 16	9.88	3.40	16,912
Level of schooling head of household when father was 16	2.98	1.47	14,614
Level of schooling head of household when mother was 16	2.99	1.46	16,912

Table 1.2: Mean schooling child by schooling level mother/father (test of MTR-MTS assumption)

Schooling level parent	Mothers		Fathers	
	$E[y z = u]^a$	$P(z=u)$	$E[y z = u]^a$	$P(z=u)$
1: Less than high school (<12 years)	12.96	0.035	13.00	0.082
2: High school (12 years)	13.98	0.627	13.85	0.495
3: Some college (13-15 years)	15.19	0.158	14.85	0.142
4: Bachelor's degree (16 years)	15.93	0.131	15.59	0.141
5: Master's degree (17 years)	16.09	0.020	15.91	0.032
6: More than a Master's degree (>17 years)	16.29	0.028	16.36	0.107
<i>N</i>	21,545		21,545	

^aMTR-MTS assumption not rejected

1.4 Results

In the analysis below we will compare the results of the nonparametric bounds analysis with the results of using an exogenous treatment selection assumption (ETS). The exogenous treatment selection assumption implies that $E[y(t)|z \neq t] = E[y|z = t]$ and yields point identification. It assumes that the schooling level of fathers and mothers is unrelated to unobserved factors affecting child's schooling (like child rearing talents or heritable endowments). Exogenous treatment selection is also assumed when regressing child's years of schooling on years of schooling of his parents. We will however not assume a linear effect of the years of schooling of the parent but instead estimate the effect of moving from one level of parental schooling to the next. Therefore we will compare the results of the bounds analysis with the results of an ETS assumption, which is the same as running OLS on child's schooling with one dummy variable for each level of mother's (father's) schooling.

Figures 1.1 and 1.2 show nonparametric bounds on mean years of schooling as a function of mother's (father's) level of schooling, compared to the exogenous treatment selection assumption. The no-assumption bounds as well as the MTR bounds are quite wide and do not give much information.⁸ In the top-right panels the monotone treatment selection assumption is added to get the MTR-MTS bounds. As was already stated in Section 1.2 this combined MTR-MTS assumption can be tested as $E[y|z = u]$ must be weakly increasing in u . Table 1.2 shows that the MTR-MTS assumption is not rejected as average years of child's schooling is indeed weakly increasing both in the level of mother's schooling as in the level of father's schooling.

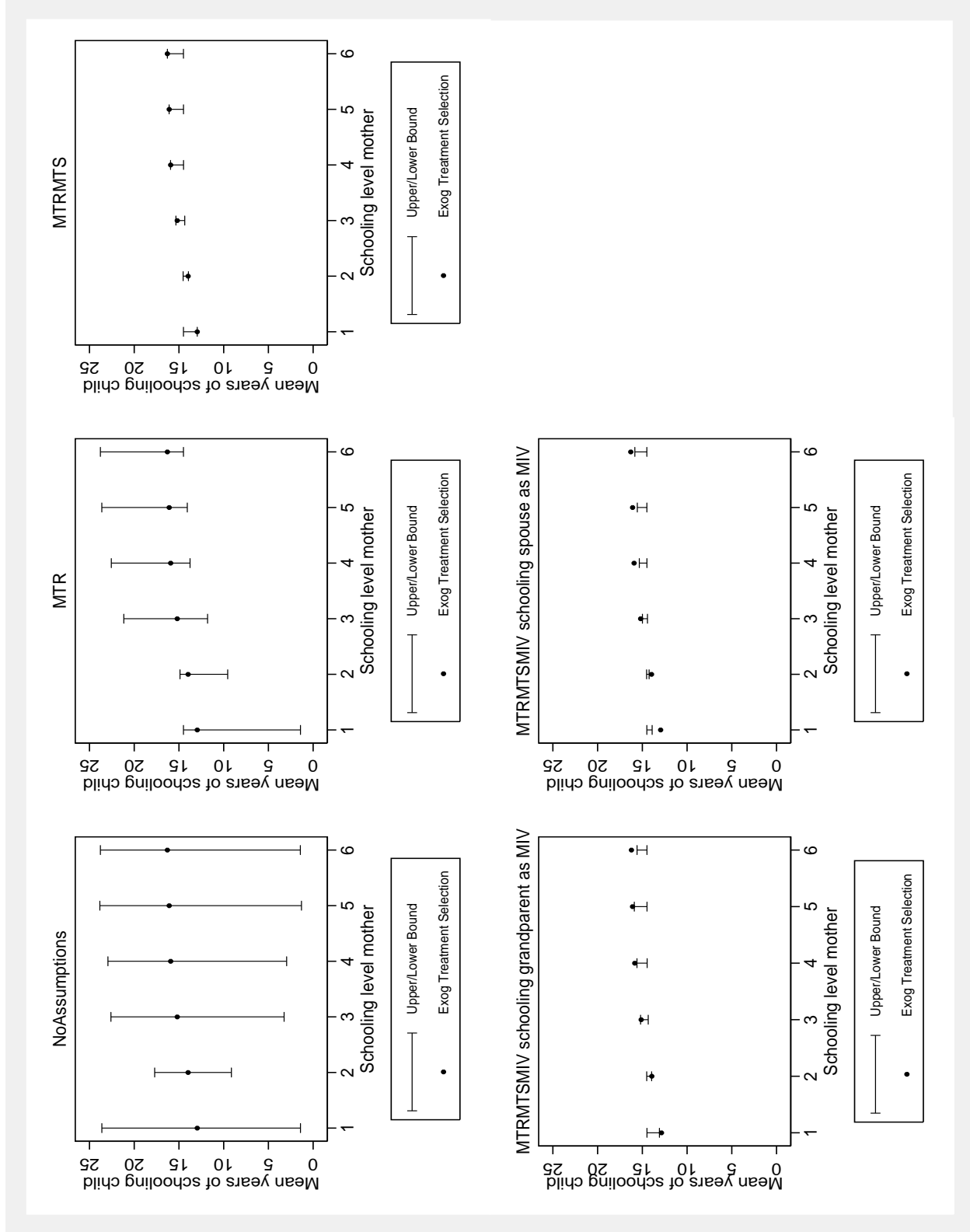
Adding the monotone treatment selection assumption strongly reduces the width of the bounds as is shown in the top-right panels of Figures 1.1 and 1.2. For the lowest levels of mother's and father's schooling the exogenous treatment selection point estimates almost coincide with the lower bounds, while for the highest levels they almost coincide with the upper bounds, but the ETS results never fall outside the MTR-MTS bounds.

In the bottom two panels of Figures 1.1 and 1.2 we add the MIV assumption. The bottom-left panels show the bounds using grandparent's schooling as a monotone instrumental variable and the bottom-right panels use the schooling of the spouse as a MIV.

Using the schooling of the grandparent as MIV gives bounds which are tighter than the MTR-MTS bounds. Both for mothers as for fathers the ETS results seem to fall outside the bounds for the highest and lowest levels of parents' schooling. The identifying power of the schooling of the spouse seems even stronger. Using spousal schooling as a MIV again reduces the bounds compared to the MTR-MTS bounds and now the point estimates fall outside the bounds for all levels of mother's schooling and for fathers this is true for the lowest and highest levels of schooling.

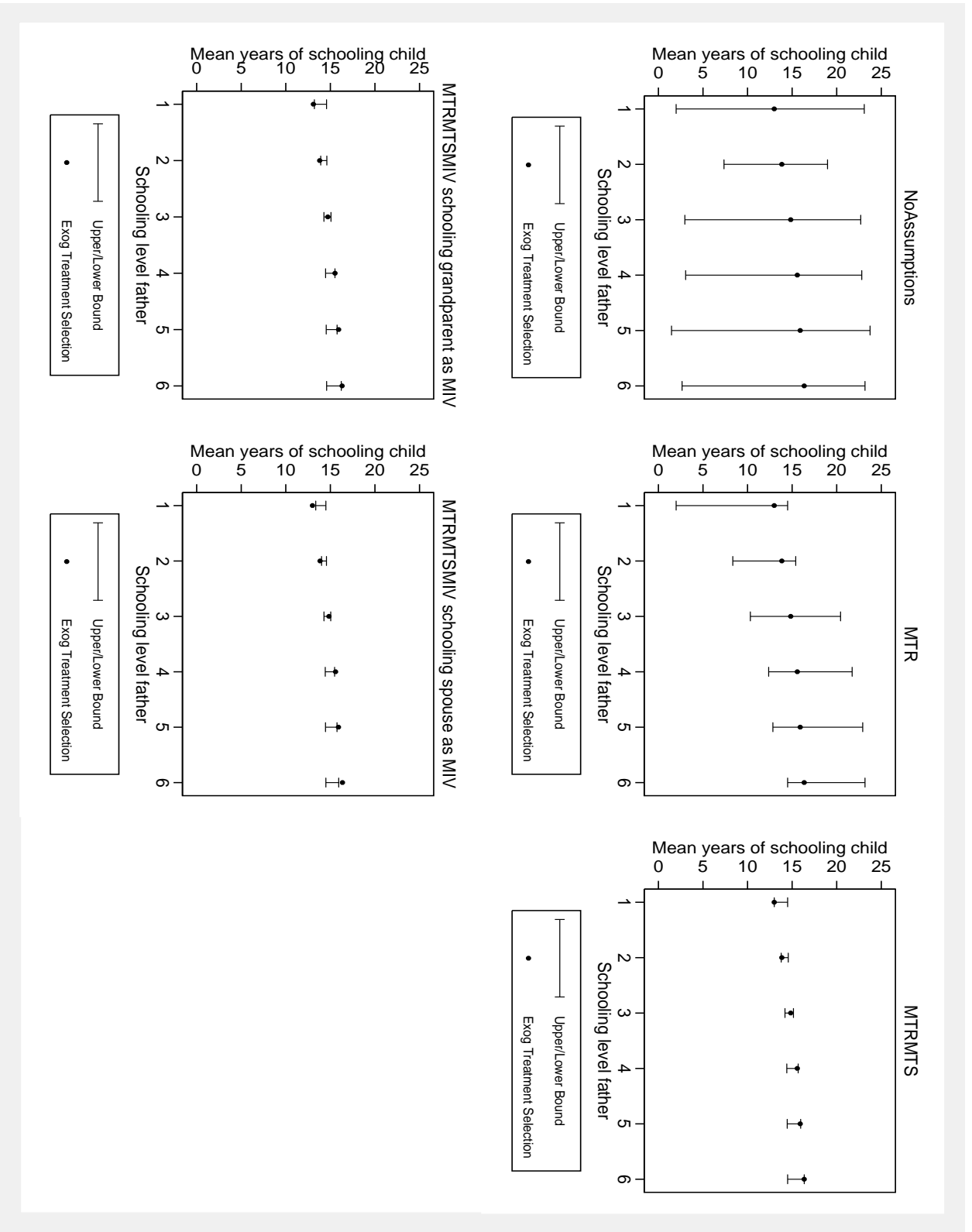
⁸For the no-assumption bounds and the MTR-bounds we take the lowest years of schooling of the child observed in the data (1 year) as \underline{y} and the highest observed years (24 years) as \bar{y} .

Figure 1.1: Child's mean schooling as function of mother's schooling; nonparametric bounds compared to ETS



Schooling levels: 1: Less than high school, 2: High school, 3: Some college, 4: Bachelor's degree, 5: Master's degree, 6: More than a Master's degree. Sample using schooling grandparent as MIV is smaller (N=16912)

Figure 1.2: Child's mean schooling as function of father's schooling: nonparametric bounds compared to ETS



Schooling levels: 1: Less than high school, 2: High school, 3: Some college, 4: Bachelor's degree, 5: Master's degree, 6: More than a Master's degree. Sample using schooling grandparent as MIV is smaller (N=14614)

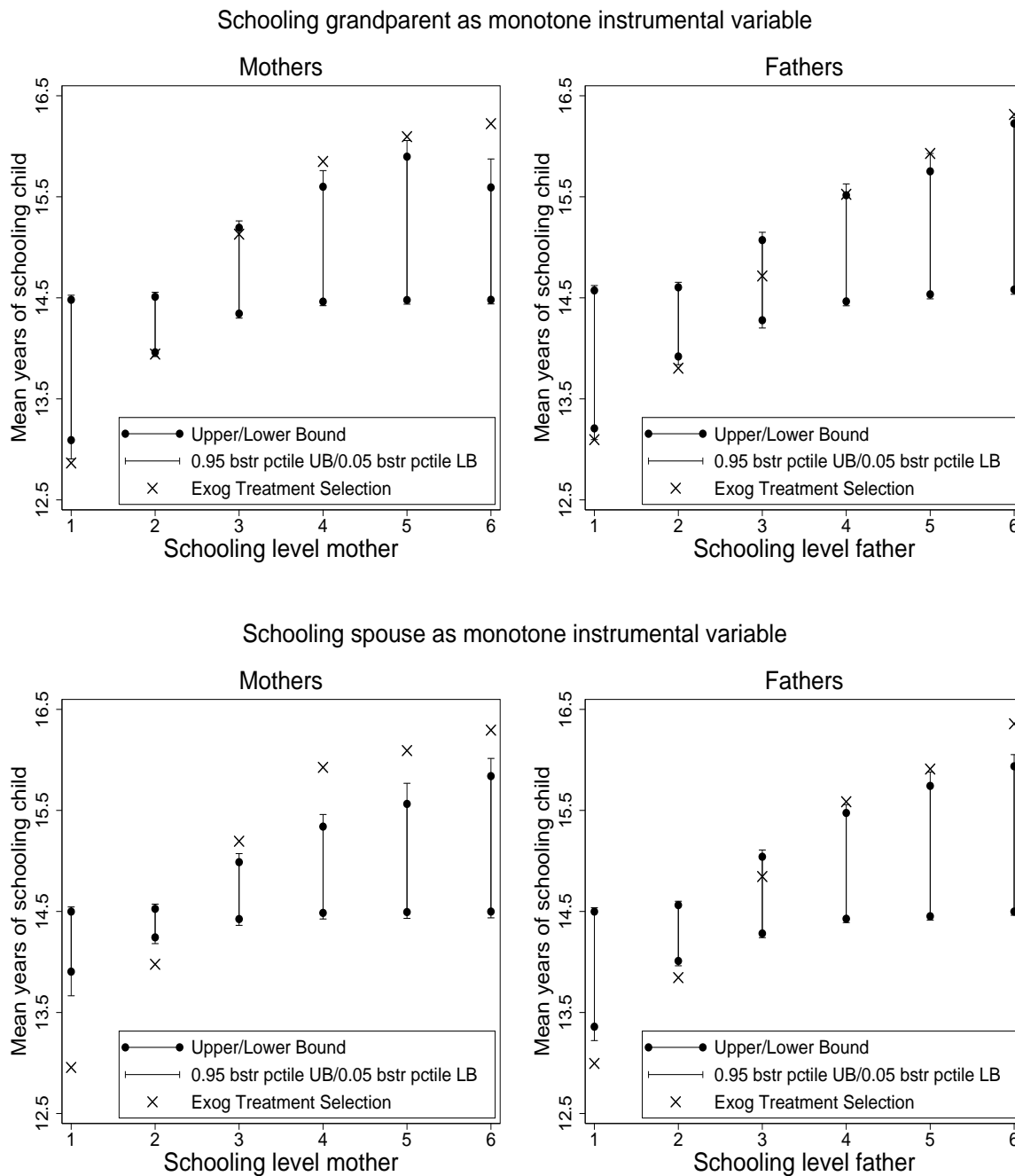
Figures 1.1 and 1.2 only show the bounds, to investigate whether the ETS results are significantly outside the bounds we will take a closer look at the MTR-MTS-MIV bounds in Figure 1.3 where we add 0.05 and 0.95 bootstrapped percentiles around the bounds.⁹ Figure 1.3 shows that when we use grandparent's schooling as MIV, the ETS estimates of the effect of mother's schooling are significantly higher than the nonparametric upper bounds for schooling levels 4 (Bachelor's degree) and 6 (more than a Master's degree) and they are just outside the confidence intervals for levels 1 and 5. For fathers some of the ETS results are just outside the bootstrapped confidence intervals around the bounds but not as much as the results for mothers.

Figure 1.3 also shows that the identifying power of the schooling of the spouse as a monotone instrumental variable is indeed stronger than grandparent's schooling. For mothers all the ETS estimates fall outside the bootstrapped confidence intervals around the bounds. Also for fathers the ETS results for schooling levels 1, 2 and 6 are significantly outside the MTR-MTS-MIV bounds. ETS underestimates for low levels and overestimates for high levels of parents' schooling.

Up to now we have only looked at bounds on $E[y(t)]$ while we are interested in the effect of increasing parents' schooling from one level to the next. Table 1.3 shows bounds on $\Delta(s, t) = E[y(t)] - E[y(s)]$ for mother's level of schooling and Table 1.4 shows the results for father's level of schooling. The ETS results range from an increase of 0.17 years of schooling when increasing mother's schooling from a bachelor's degree to a master's degree ($\Delta(4, 5)$), to an increase of 1.22 years when increasing mother's schooling from high school to some college ($\Delta(2, 3)$). The ETS results on father's schooling seem to be more constant, although the results also vary, from 0.32 years for $\Delta(4, 5)$ to 1 year for $\Delta(2, 3)$. The no-assumption bounds and the MTR bounds are again not very informative, since they are relatively wide.

⁹Bootstrapped confidence intervals are based on 1000 replications. To control for the fact that there are multiple observations from one family, the sample obtained in each replication is a bootstrap sample of clusters.

Figure 1.3: MTR-MTS-MIV bounds compared to ETS—a closer look



The 0.05 and 0.95 bootstrap percentiles are based on 1000 replications. To adjust for the fact that the sample contains multiple children from one family the sample drawn during each replication is a bootstrap sample of clusters. Levels of schooling 1: Less than high school, 2: High school, 3: Some college, 4: Bachelor's degree, 5: Master's degree, 6: More than a Master's degree. Sample using schooling grandparent as MIV is smaller (N=16912 for mothers, N=14614 for fathers)

Table 1.3: Nonparametric bounds on the effect of mother's schooling on child's years of schooling

	ETS			No-assumption bounds			
	β	95% bstr. conf. int.		0.05 pctile	LB	UB	0.95 pctile
$\Delta(1,2)$	1.02	0.80	1.23	-14.62	-14.48	16.30	16.42
$\Delta(2,3)$	1.22	1.10	1.34	-14.57	-14.47	13.47	13.58
$\Delta(3,4)$	0.73	0.60	0.86	-19.75	-19.64	19.69	19.80
$\Delta(4,5)$	0.17	-0.07	0.41	-21.70	-21.63	20.88	20.98
$\Delta(5,6)$	0.20	-0.10	0.51	-22.46	-22.41	22.48	22.52
$\Delta(2,4)$	1.95	1.85	2.05	-14.86	-14.75	13.80	13.92

	MTR bounds				MTR-MTS bounds			
	0.05 pctile	LB	UB	0.95 pctile	0.05 pctile	LB	UB	0.95 pctile
$\Delta(1,2)$	0	0	13.47	13.51	0	0	1.58	1.76
$\Delta(2,3)$	0	0	11.61	11.67	0	0	1.40	1.48
$\Delta(3,4)$	0	0	10.76	10.81	0	0	1.59	1.67
$\Delta(4,5)$	0	0	9.86	9.90	0	0	1.61	1.81
$\Delta(5,6)$	0	0	9.72	9.76	0	0	1.80	1.98
$\Delta(2,4)$	0	0	13.01	13.12	0	0	2.00	2.08

	MTR-MTS-MIV bounds grandparent as MIV ^a				MTR-MTS-MIV bounds spouse as MIV			
	0.05 pctile	LB	UB	0.95 pctile	0.05 pctile	LB	UB	0.95 pctile
$\Delta(1,2)$	0	0	1.42	1.62	0	0	0.62	0.86
$\Delta(2,3)$	0	0	1.23	1.30	0	0	0.74	0.86
$\Delta(3,4)$	0	0	1.26	1.41	0	0	0.91	1.05
$\Delta(4,5)$	0	0	1.43	1.61	0	0	1.08	1.29
$\Delta(5,6)$	0	0	1.11	1.41	0	0	1.35	1.52
$\Delta(2,4)$	0	0	1.64	1.79	0	0	1.10	1.23

The 0.05 and 0.95 bootstrap percentiles are based on 1000 replications. To adjust for the fact that the sample contains multiple children from one family, the sample drawn during each replication is a bootstrap sample of clusters. 1: Less than high school, 2: High school, 3: Some college, 4: Bachelor's degree, 5: Master's degree, 6: More than a Master's degree. ^aSample using schooling grandparent as MIV is smaller (N=16912)

Table 1.4: Nonparametric bounds on the effect of father's schooling on child's years of schooling

	ETS			No-assumption bounds			
	β	95% bstr. conf. int.		0.05 pctile	LB	UB	0.95 pctile
$\Delta(1,2)$	0.85	0.71	0.99	-15.87	-15.73	16.98	17.10
$\Delta(2,3)$	1.00	0.88	1.11	-16.12	-16.00	15.34	15.47
$\Delta(3,4)$	0.74	0.61	0.87	-19.76	-19.64	19.85	19.95
$\Delta(4,5)$	0.32	0.12	0.54	-21.42	-21.34	20.67	20.79
$\Delta(5,6)$	0.45	0.25	0.66	-21.18	-21.09	21.70	21.77
$\Delta(2,4)$	1.74	1.63	1.84	-16.04	-15.92	15.46	15.58

	MTR bounds				MTR-MTS bounds			
	0.05 pctile	LB	UB	0.95 pctile	0.05 pctile	LB	UB	0.95 pctile
$\Delta(1,2)$	0	0	13.42	13.46	0	0	1.57	1.69
$\Delta(2,3)$	0	0	12.09	12.14	0	0	1.37	1.45
$\Delta(3,4)$	0	0	11.42	11.47	0	0	1.48	1.56
$\Delta(4,5)$	0	0	10.55	10.60	0	0	1.55	1.70
$\Delta(5,6)$	0	0	10.33	10.38	0	0	1.91	1.99
$\Delta(2,4)$	0	0	13.39	13.48	0	0	1.90	1.98

	MTR-MTS-MIV bounds grandparent as MIV ^a				MTR-MTS-MIV bounds spouse as MIV			
	0.05 pctile	LB	UB	0.95 pctile	0.05 pctile	LB	UB	0.95 pctile
$\Delta(1,2)$	0	0	1.40	1.50	0	0	1.20	1.35
$\Delta(2,3)$	0	0	1.15	1.26	0	0	1.03	1.11
$\Delta(3,4)$	0	0	1.24	1.37	0	0	1.19	1.28
$\Delta(4,5)$	0	0	1.29	1.46	0	0	1.32	1.45
$\Delta(5,6)$	0	0	1.69	1.75	0	0	1.48	1.60
$\Delta(2,4)$	0	0	1.60	1.73	0	0	1.46	1.56

The 0.05 and 0.95 bootstrap percentiles are based on 1000 replications. To adjust for the fact that the sample contains multiple children from one family, the sample drawn during each replication is a bootstrap sample of clusters. 1: Less than high school, 2: High school, 3: Some college, 4: Bachelor's degree 5: Master's degree, 6: More than a Master's degree.

^aSample using schooling grandparent as MIV is smaller (N=14614)

Adding the monotone treatment selection assumption tightens the bounds significantly. It gives bounds on the treatment effects ranging from an effect between 0 and 1.40 years when increasing mother's schooling from high school to some college, to an effect between 0 and 1.80 years when increasing mother's schooling from a master's degree to more than a master's degree. For the same increases in father's schooling the effects are respectively within $[0, 1.37]$ years and within $[0, 1.91]$ years. Still these bounds include a zero effect as well as an effect as large as the ETS results and are therefore not very instructive.

When we use grandparent's schooling as a monotone instrumental variable we get upper bounds that are lower than the MTR-MTS bounds, but they are still higher than the ETS results. Using the schooling of the spouse as a MIV gives bounds which are more informative. The ETS results on father's schooling are still within the bounds, but for the effect of increasing mother's schooling from high school to some college ($\Delta (2,3)$) the ETS result falls outside the bootstrapped confidence interval around the MTR-MTS-MIV bounds.

Increasing parents' schooling from high school to a bachelor's degree

Since most fathers and mothers either have a high school degree or a bachelor's degree we can, instead of looking at the effect of moving from one level of education to the next, investigate the effect of increasing mother's/father's schooling from a high school degree (12 years) to a bachelor's degree (16 years) ($\Delta (2,4)$). Table 1.3 shows that under the exogenous treatment selection assumption the effect of increasing mother's schooling from high school to a bachelor's degree increases child's schooling on average by 1.95 years. This ETS estimate falls within the no-assumption, MTR and MTR-MTS bounds. If we however use grandparent's schooling or the schooling of the spouse as MIV we obtain upper bounds which are significantly lower than 1.95. When we use grandparent's schooling as MIV we obtain an upper bound of 1.64, and when we use the schooling of

the spouse as a monotone instrumental variable we obtain an upper bound of 1.10 years which is almost half the ETS estimate.

A similar pattern is observed when we look at the effect of increasing father's schooling from high school to a bachelor's degree in Table 1.4. The ETS estimate of this treatment effect is equal to 1.74 years. This is not significantly different from the upper bound using grandparent's schooling as MIV but it is significantly larger than the MTR-MTS-MIV upper bound of 1.46 years, when we use the schooling of the spouse as MIV.

Although the bounds do not exclude a zero effect of parents' schooling on years of schooling of the child, the upper bounds are informative since they are significantly smaller than the results obtained under the exogenous treatment selection assumption. These results are in line with the studies using twins, adoptees and some of the instrumental variables studies in the sense that most of these studies also find that OLS (ETS) overestimates the effect of parental schooling on child's schooling.

The effect of parents' schooling on the probability of a bachelor's degree

Instead of estimating the effect of parents' level of schooling on child's years of schooling we now focus on the effect on the probability that the child has a bachelor's degree (≥ 16 years of schooling). Due to compulsory schooling laws most children finish high school. The most important difference in schooling outcomes between children is the difference between having completed college or not. Tables 1.6 and 1.7 therefore show the nonparametric bounds compared to the ETS estimates of the effect of parents' schooling on the probability that a child completes college.¹⁰ The results are very similar to the results on years of schooling. The no-assumption and MTR bounds are again relatively wide. Table 1.5 shows that the MTR-MTS assumption is not rejected since the probability that the child has a bachelor's degree weakly increases with mother's and father's level of schooling. Adding the MTS assumption tightens the bounds and the bounds are narrowest when

¹⁰Since the probability of a bachelor's degree is between zero and one by definition, we take 0 as \underline{y} and 1 as \bar{y} when obtaining the no-assumption bounds and the MTR-bounds.

we add a MIV assumption, whereby the decline in the upper bound is strongest when we use the schooling of the spouse as a monotone instrumental variable.

Table 1.5: Probability that child has bachelor's degree by level of schooling mother/father

Schooling level parent	Mothers		Fathers	
	$E [P(y \geq 16yrs) z = u]^a$	$P(z=u)$	$E [P(y \geq 16yrs) z = u]^a$	$P(z=u)$
1: Less than high school (<12 years)	0.17	0.035	0.17	0.082
2: High school (12 years)	0.36	0.627	0.32	0.495
3: Some college (13-15 years)	0.58	0.158	0.52	0.142
4: Bachelor's degree (16 years)	0.75	0.131	0.70	0.141
5: Master's degree (17 years)	0.76	0.020	0.74	0.032
6: More than a Master's degree (>17 years)	0.79	0.028	0.81	0.107
<i>N</i>	21,545		21,545	

^aMTR-MTS assumption not rejected

The ETS results indicate that increasing mother's schooling from a high school degree to a bachelor's degree ($\Delta(2,4)$) increases the probability that a child has a bachelor's degree with 40 percentage points which is very similar to the effect of father's schooling of 37 percentage points. These estimates are within the no-assumption, MTR and MTR-MTS bounds. Adding a MIV assumption gives a very different picture though, since the MTR-MTS-MIV upper bounds are notably smaller than the ETS estimates. Using grandparent's schooling as MIV gives an upper bound for mothers of 33 percentage points which is significantly smaller than the ETS estimate, and for fathers the upper bound is equal 34 percentage points. The MTR-MTS-MIV bounds using the schooling of the spouse as MIV give upper bounds which are even smaller; for mothers the effect is at most 22 percentage points and for fathers at most 31 percentage points. Both upper bounds are significantly different from the ETS estimates.

Table 1.6: Nonparametric bounds on effect of mother's schooling on child's probability of a bachelor's degree

	ETS			No-assumption bounds				
	β	95% bstr. conf. int.		0.05 pctile	LB	UB	0.95 pctile	
$\Delta(1,2)^b$	0.19	0.15	0.23	-0.76	-0.75	0.59	0.60	
$\Delta(2,3)$	0.23	0.20	0.25	-0.51	-0.50	0.71	0.72	
$\Delta(3,4)$	0.17	0.14	0.20	-0.84	-0.83	0.88	0.88	
$\Delta(4,5)$	0.01	-0.05	0.06	-0.96	-0.95	0.90	0.90	
$\Delta(5,6)$	0.03	-0.03	0.09	-0.98	-0.97	0.98	0.98	
$\Delta(2,4)$	0.40	0.38	0.42	-0.50	-0.50	0.74	0.75	
	MTR bounds				MTR-MTS bounds			
	0.05 pctile	LB	UB	0.95 pctile	0.05 pctile	LB	UB	0.95 pctile
$\Delta(1,2)$	0	0	0.48	0.49	0	0	0.30	0.33
$\Delta(2,3)$	0	0	0.66	0.67	0	0	0.26	0.28
$\Delta(3,4)$	0	0	0.64	0.64	0	0	0.33	0.35
$\Delta(4,5)$	0	0	0.57	0.58	0	0	0.30	0.35
$\Delta(5,6)$	0	0	0.56	0.57	0	0	0.33	0.37
$\Delta(2,4)$	0	0	0.73	0.74	0	0	0.41	0.42
	MTR-MTS-MIV bounds grandparent as MIV ^a				MTR-MTS-MIV bounds spouse as MIV			
	0.05 pctile	LB	UB	0.95 pctile	0.05 pctile	LB	UB	0.95 pctile
$\Delta(1,2)$	0	0	0.27	0.30	0	0	0.12	0.19
$\Delta(2,3)$	0	0	0.23	0.25	0	0	0.13	0.15
$\Delta(3,4)$	0	0	0.26	0.30	0	0	0.19	0.22
$\Delta(4,5)$	0	0	0.24	0.29	0	0	0.17	0.23
$\Delta(5,6)$	0	0	0.19	0.26	0	0	0.26	0.29
$\Delta(2,4)$	0	0	0.33	0.37	0	0	0.22	0.25

The 0.05 and 0.95 bootstrap percentiles are based on 1000 replications. To adjust for the fact that the sample contains multiple children from one family, the sample drawn during each replication is a bootstrap sample of clusters. 1: Less than high school, 2: High school, 3: Some college, 4: Bachelor's degree 5: Master's degree, 6: More than a Master's degree.

^aSample using schooling grandparent as MIV is smaller (N=16912)

Table 1.7: Nonparametric bounds on effect of father's schooling on child's probability of a bachelor's degree

	ETS			No-assumption bounds			
	β	95% bstr. conf. int.		0.05 pctile	LB	UB	0.95 pctile
$\Delta(1,2)$	0.15	0.12	0.17	-0.78	-0.77	0.65	0.66
$\Delta(2,3)$	0.20	0.17	0.23	-0.60	-0.59	0.77	0.78
$\Delta(3,4)$	0.17	0.14	0.20	-0.84	-0.83	0.88	0.89
$\Delta(4,5)$	0.04	-0.00	0.09	-0.94	-0.93	0.89	0.90
$\Delta(5,6)$	0.07	0.03	0.12	-0.91	-0.90	0.96	0.96
$\Delta(2,4)$	0.37	0.35	0.40	-0.57	-0.57	0.80	0.80

	MTR bounds				MTR-MTS bounds			
	0.05 pctile	LB	UB	0.95 pctile	0.05 pctile	LB	UB	0.95 pctile
$\Delta(1,2)$	0	0	0.51	0.52	0	0	0.30	0.32
$\Delta(2,3)$	0	0	0.69	0.69	0	0	0.28	0.29
$\Delta(3,4)$	0	0	0.68	0.69	0	0	0.32	0.33
$\Delta(4,5)$	0	0	0.63	0.63	0	0	0.31	0.34
$\Delta(5,6)$	0	0	0.61	0.62	0	0	0.36	0.38
$\Delta(2,4)$	0	0	0.76	0.76	0	0	0.40	0.42

	MTR-MTS-MIV bounds grandparent as MIV ^a				MTR-MTS-MIV bounds spouse as MIV			
	0.05 pctile	LB	UB	0.95 pctile	0.05 pctile	LB	UB	0.95 pctile
$\Delta(1,2)$	0	0	0.26	0.29	0	0	0.24	0.27
$\Delta(2,3)$	0	0	0.23	0.25	0	0	0.21	0.22
$\Delta(3,4)$	0	0	0.27	0.29	0	0	0.26	0.28
$\Delta(4,5)$	0	0	0.25	0.29	0	0	0.26	0.30
$\Delta(5,6)$	0	0	0.31	0.33	0	0	0.30	0.32
$\Delta(2,4)$	0	0	0.34	0.36	0	0	0.31	0.34

The 0.05 and 0.95 bootstrap percentiles are based on 1000 replications. To adjust for the fact that the sample contains multiple children from one family, the sample drawn during each replication is a bootstrap sample of clusters. 1: Less than high school, 2: High school, 3: Some college, 4: Bachelor's degree, 5: Master's degree, 6: More than a Master's degree.^aSample using schooling grandparent as MIV is smaller (N=14614)

Figure 1.4: Bounds on the effect of increasing parents' schooling from high school to a bachelor's degree

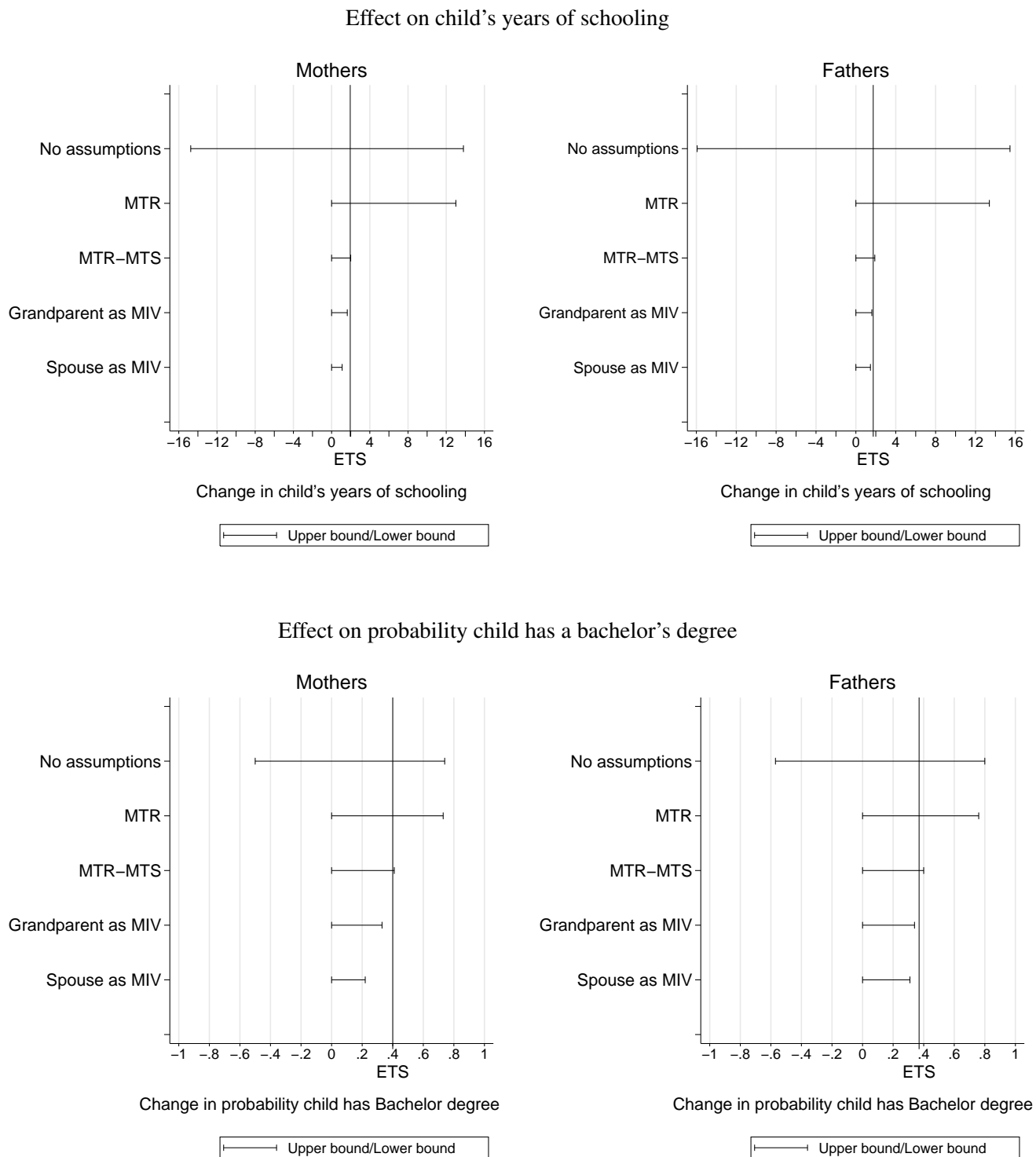


Figure 1.4 shows the bounds on the treatment effects of increasing mother's (father's) schooling from a high school degree to a bachelor's degree on child's years of schooling (top panel) and on the probability that the child has a bachelor's degree (bottom panel). These pictures clearly show how the bounds are tightened by adding the MTR, MTS and MIV assumptions. Both for the effect on years of schooling as the for effect on the probability of a bachelor's degree the ETS results clearly overestimate the effect of parents' schooling compared to the MTR-MTS-MIV upper bounds. This effect is strongest for mothers, since here the upper bounds (using the schooling of the spouse as MIV) are almost half the ETS results.

1.5 Conclusion

Regressing child's schooling on parents' schooling generally gives large positive and significant estimates. Since these estimates need not be equal to the true causal relation, different identification strategies have been used. These identification approaches generally put strong requirements on the data since you need a large data set with completed schooling outcomes of both parents and their children and you either need a large sample of twins or adoptees or a good instrument. And even if you are able to apply any of these identification strategies, you will always have to make a number of assumptions in order to interpret the results as causal.

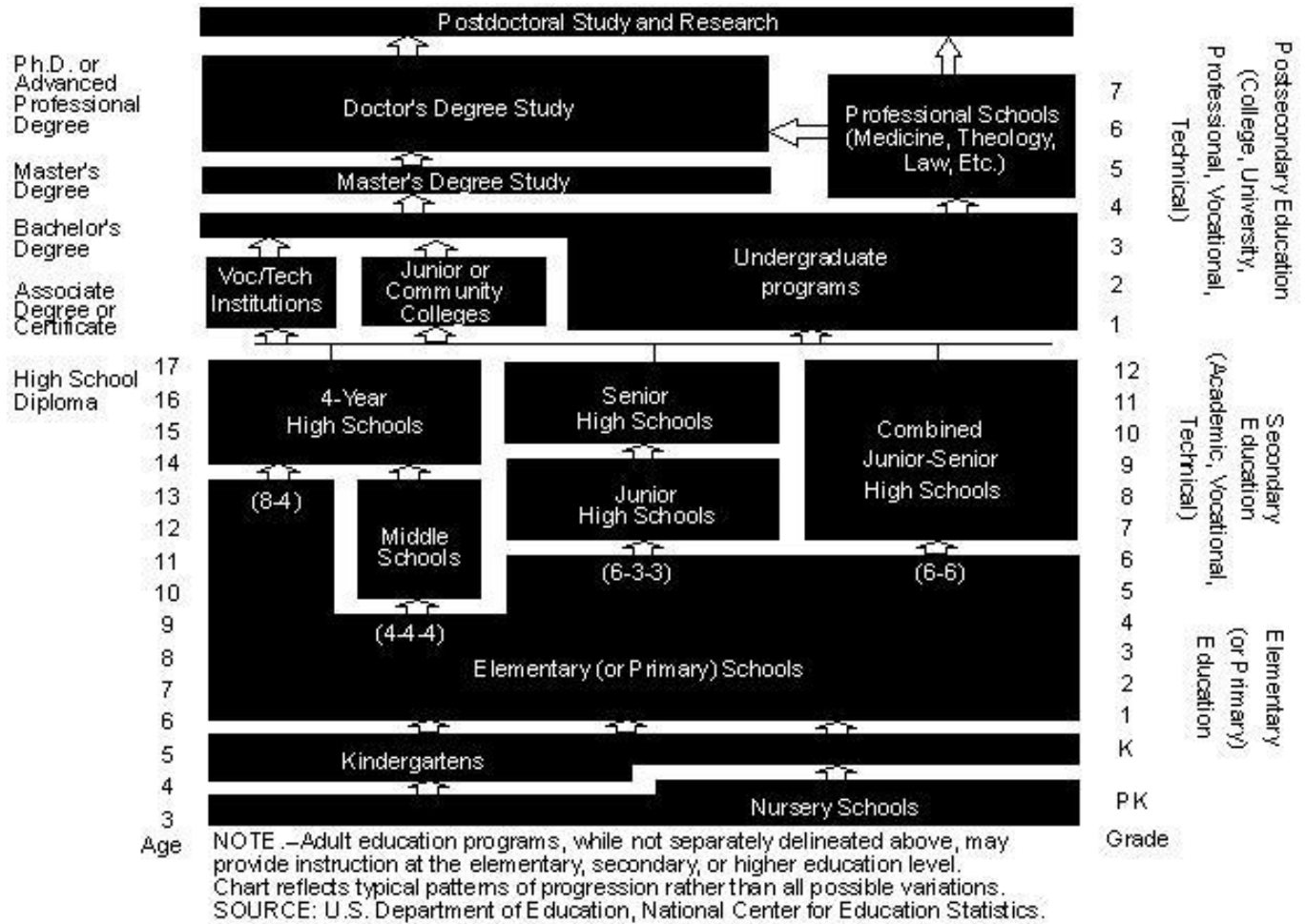
This chapter used a relatively new approach to learn more about the effect of father's and mother's schooling on the schooling of their child. By making relatively weak and testable assumptions we have obtained bounds on the effect of increasing parents' schooling on years of schooling of the child and on the probability that the child obtains a bachelor's degree. We started with obtaining bounds without making any assumptions and then tightened the bounds by subsequently adding a monotone treatment response assumption (MTR), a monotone treatment selection assumption (MTS) and a monotone instrumental

variable assumption (MIV), whereby we used the schooling of the grandparent and the schooling of the spouse as monotone instrumental variables.

Although the bounds on the treatment effects include a zero effect, the upper bounds are informative especially for the effect of increasing parents' schooling from a high school degree to a bachelor's degree. For mothers the MTR-MTS-MIV upper bounds are almost half the estimates under the exogenous treatment selection assumption. Also for the effect of increasing father's schooling from high school to a bachelor's degree the estimates under the exogenous treatment selection assumption are significantly larger than the MTR-MTS-MIV upper bounds. The results in this chapter show that the effect of parents' schooling is lower than what one would conclude on the basis of simple correlations and that there might even be no effect at all. These findings are in line with the studies using twins, adoptees and some of the instrumental variables studies in the sense that most of these studies also find that OLS (ETS) overestimates the effect of parental schooling on child's schooling.

Appendix

Figure 1.5: Map of the U.S. Education System



(<http://www.ed.gov/about/offices/list/ous/international/usnei/us/edlite-map.html>)

Chapter 2

Estimating intergenerational schooling mobility on censored samples: Consequences and remedies¹

2.1 Introduction

Most empirical studies on intergenerational mobility estimate a version of the following model

$$Y_t = \alpha + \beta Y_{t-1} + u_t \quad (2.1)$$

where t is a generation index, Y_t and Y_{t-1} represent *realized* outcomes of child and parent, and u_t is a child-specific characteristic. In most studies the parameter of interest is β which measures the outcome association between parent and child. Estimating β , however, puts strong requirements on data. In household surveys, in particular, the collection of information on realized outcomes of children is often problematic. If Y represents (permanent) income, for example, income information of children is rarely available. Most children who still live with their parents do not work. And even if information is available, intergenerational mobility estimates will be biased downwards when the children's income is measured too early in life (Haider and Solon (2006); B lmark and Lindquist (2006)).

¹Joint work with Erik Plug

In this chapter we let Y be years of schooling, and focus on the problem that (some) children may still be in school at the time of data collection, which goes under name of the censoring problem. We consider this a serious problem for three reasons. First, we cannot ignore censoring empirically, because if we do, least squares regression on censored samples would give us intergenerational mobility estimates that are too low. Second, we observe that censoring is a widely spread phenomenon. Of the recent studies that (aim to) make a distinction between causation and selection, almost all rely on samples with incomplete information on adult children. Among these studies are Behrman and Rosenzweig (2002); Chevalier (2004); Currie and Moretti (2003); Plug (2004); Black *et al.* (2005A); Carneiro *et al.* (2007); Oreopoulos *et al.* (2006); Maurin and McNally (2008). And third, the solutions offered to handle censored samples rely on assumptions that may not hold in practice, resulting in biased mobility estimates.

The natural solution to the censoring problem is patience. If researchers were patient and could wait until all children in the censored sample finished their schooling to collect their data, we wouldn't need to worry about censoring. Unfortunately, many researchers tend to be impatient. They are, presumably, more interested in the degree of intergenerational mobility among current generations than previous generations and are therefore willing to estimate parental schooling effects on censored samples using correction methods that do not always work. Since the latter approach certainly merits serious consideration, it is important to know (more) about how the available correction methods deal with censored observations.

Three correction methods are currently in use: maximum likelihood approach, replacement of observed with expected years of schooling, and elimination of all school-aged children.² In this chapter we apply these three different methods to correct for the censor-

²An alternative method to deal with censored observations is to look at intermediate outcomes that are realized and available, such as birth weight (Currie and Moretti 2004), grade repetition (Carneiro *et al.* 2007; Oreopoulos *et al.* 2006; Maurin and McNally 2008) or post-compulsory schooling attendance (Chevalier 2004). Without information on realized school outcomes of children, however, we do not know how informative these intermediate outcomes are when it comes to assessing intergenerational schooling effects.

ing problem to one particular data set: the most recent version of the Wisconsin Longitudinal Study (henceforth often WLS).

The WLS collects information on a large group of students who graduated from Wisconsin high schools in 1957. In 1975, 1992 and 2004 the same students were contacted again and asked about their children's schooling. The questions cover three different school stages. In 1975 most children are still in school: the sample includes information on expected schooling. In 1992 the same children are about to complete school: the sample is a censored sample. In 2004 all children have passed their school-going age: the sample contains information on completed schooling.

Our contributions are twofold. First, we present new estimates of the intergenerational mobility of schooling. In addition, we make a distinction between own birth children and adoptees to investigate how much inherited abilities contribute to the impact of parental schooling. With updated 2004 samples, we estimate the ultimate mobility models in which censored observations are absent. And second, we examine the validity of the different solutions to deal with the problem of censored data. With the 1975 and 1992 samples, we estimate the impact of parental schooling on children's schooling applying the various procedures to correct for censored observations and use the difference between ultimate and corrected mobility estimates as a validity indicator.

This chapter continues as follows. Section 2.2 models the intergenerational mobility of schooling, focuses on the problem that children who are still in school generate censored observations, and provides some intuition of the various solutions to it. Section 2.3 provides a brief description of the data. Section 2.4 presents and compares the parameter estimates. Section 2.5 evaluates the correction methods and presents a number of robustness tests and Section 2.6 concludes.

2.2 Mobility models using censored data

Much work on intergenerational schooling mobility has concentrated on estimating a version of the following model

$$S_t = \alpha + \beta S_{t-1} + \varepsilon_t, \quad (2.2)$$

where t is a generation index, S_t and S_{t-1} represent the schooling of child and parent, usually measured as the number of years of completed schooling, and ε_t is a child-specific characteristic. The mobility parameter β measures the association between the schooling of parent and child. With information on S_t and S_{t-1} , the least-squares estimator is defined as

$$\text{plim } \beta_{LS} = \text{cov}(S_t, S_{t-1}) / \text{var}(S_{t-1}) = \beta. \quad (2.3)$$

A well-known problem in analyzing intergenerational schooling mobility is that information on the child's completed schooling is not always available. Some children are still in school at the time data are collected and create censored observations. To accommodate censored observations, the intergenerational schooling model needs to be rewritten as

$$S_t^c = \begin{cases} S_t^c = S_t & \text{if } d_t = 0, \\ S_t^c < S_t & \text{if } d_t = 1, \end{cases} \quad (2.4)$$

where S_t^c represents the child's years of schooling observed in the censored sample, and d_t denotes whether observations are censored ($d_t = 1$) or not ($d_t = 0$). If we would ignore censoring, and treat the children's observed years of schooling as if it were their completed years, the estimation of S_t^c on S_{t-1} using ordinary least squares gives us a mobility parameter that is too low. The intuition is as follows. We know that (a) more schooled children (with more schooled parents) are more likely to be censored; and (b) observed years are smaller than or equal to the completed years. Taken together, these observations imply that observed years of schooling covary less with parental years of schooling

$(\text{cov}(S_t^c, S_{t-1}) \leq \text{cov}(S_t, S_{t-1}))$. When we now apply least squares to estimate the model

$$S_t^c = \alpha^c + \beta^c S_{t-1} + \varepsilon_t^c, \quad (2.5)$$

it follows naturally that the corresponding least squares estimator is biased toward zero, as

$$\text{plim } \beta_{LS}^c = \text{cov}(S_t^c, S_{t-1}) / \text{var}(S_{t-1}) \leq \text{cov}(S_t, S_{t-1}) / \text{var}(S_{t-1}) = \beta. \quad (2.6)$$

Recent work on intergenerational mobility of schooling has taken three approaches to tackle the censoring problem: maximum likelihood approach, replacement of observed with expected years of schooling, and elimination of all school-aged children. Below we shortly discuss the different approaches.

A censored regression model

Plug (2004) exploits the 1992 wave of the WLS to estimate the effect of fathers and mothers schooling on child's schooling using samples of biological and adopted children. In 1992, however, many children have not yet finished their schooling (about 25% of the biological children and 40% of the adopted children). As we already mentioned, not taking censoring into account gives inconsistent estimates. Plug therefore uses a censored regression model, one of the standard procedures for handling censored observations. Assuming the conditional distribution of ε_t is normal with homoskedastic errors the likelihood function is

$$L(\theta) = \prod_{i=1}^N [\phi(S_t | S_{t-1}, \theta)]^{1-d_t} [1 - \Phi(S_t^c | S_{t-1}, \theta)]^{d_t}, \quad (2.7)$$

where ϕ and Φ represent normal density and distribution functions, θ are the distribution parameters that include β , and i indexes the family in which the child is born and raised. Maximization of (2.7) yields a consistent estimator of β , unless the error distribution is incorrectly specified, being non-normally distributed or having heteroskedastic errors of unknown form.

Eliminating all school-going aged children

Black *et al.* (2005A) estimate the effect of parental schooling on child schooling using a reform in compulsory schooling in Norway during the sixties and early seventies to draw causal inferences. Because Black *et al.* focus on relatively young parents –only those between 42 and 53 years old are affected by the reform– many children have not finished their schooling yet by the time they appear in their sample. They take account of the censoring problem by eliminating all children younger than age 25.

Many of these children are likely to have parents who were very young when they were born. Black *et al.* therefore run the risk of introducing sample selection bias when they reduce their sample. The argument is that censoring is not random but related to observed and unobserved parental characteristics, and that the corresponding estimate of the effect of parental schooling using the reduced sample can be biased.

Inserting parental expectations for children still in school

Behrman and Rosenzweig (2002) employ a mail survey –issued in 1994– to collect information on the families of identical twins born between 1936 and 1955, all drawn from the Minnesota Twin Registry (MTR). The survey contains information on the schooling of the twins, their parents and children, including information on expected schooling for children who had not completed their schooling yet; this is the case for more than 50% of their sample.³

Behrman and Rosenzweig propose to replace their censored observations with parental expectations and treat these expectations as if they were school realizations for children

³The American Economic Review provides data and programmes for replication purposes online. From this source we have extracted the twin sample using data and programmes of Antonovics and Goldberger (2005). We are able to trace 844 monozygotic twin parents with children. Of these 844 children, 428 are still in school in 1994.

with unfinished schooling. This gives the following school variable for the child

$$\tilde{S}_t = \begin{cases} S_t & \text{if } d_t = 0, \\ S_t^e & \text{if } d_t = 1, \end{cases} \quad (2.8)$$

where S_t^e represents the school level the parent expects her child to complete. Suppose we model parental expectations about their children's completed years of schooling as follows

$$S_t^e = S_t + \eta_t, \quad (2.9)$$

where η_t is the error parents make in predicting their child's completed schooling.⁴ Combining (2.2), (2.8) and (2.9) leads to

$$\tilde{S}_t = \alpha + \beta S_{t-1} + d_t \eta_t + \varepsilon_t. \quad (2.10)$$

Applying least squares to the bivariate regression of \tilde{S}_t on S_{t-1} gives us the following probability limit of the slope coefficient

$$\text{plim } \tilde{\beta}_{LS} = \text{cov}(\tilde{S}_t, S_{t-1}) / \text{var}(S_{t-1}) = \beta + \text{cov}(d_t \eta_t, S_{t-1}) / \text{var}(S_{t-1}). \quad (2.11)$$

Only if $\text{cov}(d_t \eta_t, S_{t-1})$ equals 0, Behrman and Rosenzweig's original solution produces an unbiased estimate of β . If not, the validity of the method will depend on how much the prediction error correlates with parental education and on the number of censored observations. Whether or not $\text{cov}(d_t \eta_t, S_{t-1})$ equals 0 is an empirical issue, which we will put to the test later on in this chapter.

⁴We omit subscript i here, but we do not assume that the prediction error is the same for all individuals, nor do we assume anything about the distribution of η_t .

2.3 Data

Our analysis employs the Wisconsin Longitudinal Study (WLS) of 10,317 randomly sampled graduates from Wisconsin high schools in 1957. After the initial wave of data collection, primary respondents were re-interviewed in 1975, 1992 and 2004. Together with their parents' interview of 1964, these waves provide information on, among others, educational attainment of the original graduates, their parents and children. The original sample is broadly representative for white men and women, who have completed at least twelve years of schooling. For more detailed information on the WLS we refer to Sewell *et al.* (2004) and WLS (2006) and the references therein.

In this chapter we use all three waves and exploit those questions that are targeted at the educational attainment of the respondent's children. In 1975 children are still in school and questions are asked to elicit parental expectations.⁵ In 1992 children are about to complete or just completed their schooling and information is collected on the highest grade of regular school ever attended whether the highest grade is completed or not; and whether the highest grade is obtained during the survey year. In addition, respondents are asked whether their child completed the grade or year and whether their child attended a regular school (elementary, secondary, colleges, and universities) in the past 12 months. In 2004 these children all finished their education, and respondents are asked to update their information regarding their children's completed schooling.

Our sample includes married respondents with children, who are observed in the years 1975, 1992 and 2004. In 2004 information is gathered from 7,265 of the 10,317 original respondents, of whom 5,630 are married and have children older than 12 in 1992. Of these 5,630 respondents 442 drop out because relevant schooling information of themselves, their spouses and children is missing.⁶ This leaves us with a sample of 5,188 re-

⁵Parental expectations are expressed in levels. We convert levels into years in a similar way as Antonovics and Goldberger (2005, pp.1739) recode levels into years of schooling: less than high school...10; high school graduate...12; technical and vocational education...13; some college...14; college graduate...16; M.A. or M.S. degree...18; Law degree, M.D., D.D.S., D.V.M. degree...19; Ph.D....20.

⁶For some children who finished schooling in 1992, reported years of schooling in 2004 differs from years of schooling

spondents having 14,524 own birth children and 520 adopted children. Note that in 1975 respondents are asked to express their school expectations for only one of their children. This child is randomly selected by the interviewer. The censoring analysis in which we replace censored with expected school measures relies therefore on a much smaller sample, consisting of 4,097 own birth and 52 adopted children. The selected children are on average 10 years old when the parent forms expectations regarding the child's schooling. Expectations are elicited from the originally sampled graduates which include both men and women, so the expectations are in some cases formed by the father and in other cases by the mother. Summary statistics appear in Table 2.1.

Table 2.1: Means and standard deviations of selected variables in WLS samples

	Own birth children		Adoptees	
	Mean	Std. dev.	Mean	Std. dev.
Completed years of schooling (2004)	14.37	2.28	14.03	2.09
Observed years of schooling (1992)	13.84	2.34	13.25	2.12
Expected years of schooling ^a (1975)	14.78	1.93	15.08	1.97
Years of schooling mother	12.83	1.65	13.27	1.92
Years of schooling father	13.50	2.66	14.49	2.97
Observation censored in 1992	0.23	0.42	0.39	0.49
Gender (daughter)	0.50	0.50	0.49	0.50
Age (1992)	26.52	4.51	23.97	4.62
<i>N</i>	14,524		520	

^aParental expectations are asked for only one of the respondent's children. Means and standard deviations are therefore calculated on smaller samples of respectively 4,097 and 52 observations.

reported in 1992. For these observations we replace reported schooling in 1992 and 2004 by the maximum of the two. This is done for 487 own birth children and 39 adoptees.

2.4 Results

Table 2.2 presents estimates that come from our child-parent schooling regressions run on uncensored and censored samples of own birth children and their parents. All regressions include individual controls for the child's age and gender. These parameters are not reported.⁷

In the first three columns we report estimates using the completed school measures as recorded in the 2004 sample. In columns (1) and (2) the mother's and father's schooling measures are included as separate regressors. We find that more schooled parents have more schooled children, and that more schooled mothers matter more than more schooled fathers. In column (3) the mother's and father's schooling measures are included simultaneously to control for assortative mating effects. We still find that more schooled parents get more schooled children, but that fathers and mothers now contribute equally to their offspring.

In the second three columns we estimate the same three equations using the observed school measures as recorded in the 1992 sample. With data that are partly censored we find, as expected, that all parental schooling estimates fall. It is clear that these estimates are biased. The last three columns, in which we express the difference between mobility estimates run on the censored and uncensored samples, indicate that the downward bias caused by the censoring is statistically significant and varies between the 6 and 16 percent.⁸

In the next three panels we report the estimates using alternative approaches to tackle the censoring problem: maximum likelihood approach, elimination of all school-aged children, and replacement of observed with expected years of schooling.

⁷The estimations use all children, including all children raised in one family. With multiple family observations, standard errors are not independent within families and are biased downwards. We therefore estimate the model with clustered error terms to control for correlation within families.

⁸The previous schooling models are estimated combining both WLS samples where all coefficients vary by sample status. The interacted schooling estimate represents the absolute difference between mobility parameters.

Table 2.2: Estimates of the effects of mother's and father's schooling on own birth children's schooling

	Mobility Estimates without Censoring (WLS 2004)			Mobility Estimates with Censoring (WLS 1992)			Estimated Differences		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mother's schooling	0.46*** (0.01)		0.24*** (0.02)	0.41*** (0.01)		0.20*** (0.02)	-0.05*** (0.01)		-0.04*** (0.01)
Father's schooling		0.34*** (0.01)	0.26*** (0.01)		0.31*** (0.01)	0.24*** (0.01)		-0.03*** (0.00)	-0.02*** (0.01)
<i>N</i>		14,524			14,524				
<i>N^c</i>		0			3,278				
CENSORED REGRESSION MODEL									
Mother's schooling				0.54*** (0.02)		0.29*** (0.02)	0.08*** (0.01)		0.05*** (0.01)
Father's schooling					0.38*** (0.01)	0.29*** (0.01)		0.04*** (0.00)	0.03*** (0.00)
<i>N</i>					14,524				
<i>N^c</i>					3,278				
EXCLUDING ALL CHILDREN YOUNGER THAN 25									
Mother's schooling				0.50*** (0.02)		0.26*** (0.02)	0.04*** (0.01)		0.03*** (0.01)
Father's schooling					0.36*** (0.01)	0.28*** (0.01)		0.02*** (0.01)	0.02*** (0.01)
<i>N</i>					10,143				
<i>N^c</i>					508				
CENSORED OBSERVATIONS REPLACED WITH PARENTAL EXPECTATIONS^a									
Mother's schooling	0.45*** (0.02)		0.24*** (0.02)	0.45*** (0.02)		0.23*** (0.02)	-0.00 (0.01)		-0.01 (0.01)
Father's schooling		0.35*** (0.01)	0.27*** (0.01)		0.35*** (0.01)	0.27*** (0.01)		0.00 (0.01)	0.00 (0.01)
<i>N</i>		4,097			4,097				
<i>N^c</i>		0			874				

All regressions include additional controls on the child's age and gender. Standard errors (in parentheses) allow for correlation within families. * significant at 10% level, ** significant at 5% level, *** significant at 1% level.

^aSamples are smaller because expectations are elicited for only one of the respondent's children.

We find that the corrections do not affect our results qualitatively. In all three panels the estimates reported in columns (4), (5) and (6) show that more schooled parents get more schooled children and that mothers only matter more when parental schooling estimates include assortative mating effects. But we do find that the corrections affect our results quantitatively. When compared to the uncorrected regression results using the censored sample, all three approaches remove the downward bias and give us –as they should– higher mobility estimates. When compared to those estimates obtained using the ultimate uncensored sample, the estimated differences in columns (7), (8) and (9) indicate that especially maximum likelihood and elimination approaches lead to mobility estimates that are too high. Instead of providing consistent estimates, these two censoring corrections cause an upward bias that is statistically significant and varies between the 6 and 21 percent. The medicine appears to be no better than the malady. The approach to treat parental expectations for young children as if they were realizations of completed schooling, however, does better. The bias is at most 4 percent and never statistically significant.

Adoption results

Recall that all the positive mobility estimates reported in Table 2.2 include the contribution of inherited abilities to intergenerational schooling transfers. To get rid of the effects caused by the parents' genes, we run our child-parent schooling regressions on samples of adoptees and their adoptive parents. This is done in Table 2.3 which has the same format as Table 2.2. We do this for two reasons. The first reason is that recent intergenerational mobility studies focus their attention on the parameter that measures the causal impact of the parent's schooling on that of the child. The second reason is that it is not obvious how censoring affects the schooling association between parent and child, net of the inherited endowments of parents.

Table 2.3: Estimates of the effects of mother's and father's schooling on adopted children's schooling

	Mobility Estimates without Censoring (WLS 2004)			Mobility Estimates with Censoring (WLS 1992)			Estimated Differences		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mother's schooling	0.22*** (0.05)		0.08 (0.06)	0.19*** (0.05)		0.06 (0.06)	-0.03 (0.05)		-0.02 (0.05)
Father's schooling		0.22*** (0.03)	0.19*** (0.04)		0.19*** (0.03)	0.17*** (0.03)		-0.03 (0.03)	-0.02 (0.03)
<i>N</i>		520			520				
<i>N^c</i>		0			198				
CENSORED REGRESSION MODEL									
Mother's schooling				0.21*** (0.07)		0.06 (0.07)	-0.01 (0.03)		-0.02 (0.04)
Father's schooling					0.23*** (0.04)	0.21*** (0.04)		0.01 (0.02)	0.02 (0.02)
<i>N</i>					520				
<i>N^c</i>					198				
EXCLUDING ALL CHILDREN YOUNGER THAN 25									
Mother's schooling				0.32*** (0.08)		0.19** (0.09)	0.10 (0.07)		0.11 (0.08)
Father's schooling					0.26*** (0.05)	0.20*** (0.06)		0.04 (0.05)	0.01 (0.05)
<i>N</i>					216				
<i>N^c</i>					19				
CENSORED OBSERVATIONS REPLACED WITH PARENTAL EXPECTATIONS ^d									
Mother's schooling	0.13 <i>0.15</i>		-0.13 <i>0.18</i>	0.25 <i>0.16</i>		-0.02 <i>0.19</i>	0.12 <i>0.08</i>		0.11 <i>0.10</i>
Father's schooling		0.23 <i>0.09**</i>	0.28 <i>0.11**</i>		0.28 <i>0.09***</i>	0.28 <i>0.12**</i>		0.05 <i>0.07</i>	0.00 <i>0.09</i>
<i>N</i>		52			52				
<i>N^c</i>					21				

All regressions include additional controls on the child's age and gender. Standard errors (in parentheses) allow for correlation within families. * significant at 10% level, ** significant at 5% level, *** significant at 1% level.

^dSamples are smaller because expectations are elicited for only one of the respondent's children.

In columns (1), (2) and (3) the results using the uncensored sample of adoptees are shown. We find that all the estimated effects of parental schooling drop when we move from own birth children to adoptees. This is consistent with the idea that part of the child's schooling is inherited. In column (3) we take the impact of the marriage partner into account, and find that the estimates fall only little for fathers, but much more for mothers. The maternal schooling effect reduces to 0.08 and lacks statistical significance, while the paternal schooling effect remains much larger in magnitude: 0.22 and 0.19 with or without taking into account the effect of his marriage partner. These findings are, as such, fully in line with those reported in Plug (2004) but also in Behrman and Rosenzweig (2002, 2005) and Björklund *et al.* (2006).

In columns (4), (5) and (6) of the first panel we see that the estimated effects remain qualitatively very similar, except that they are all smaller than the corresponding point estimates in the first three columns. This is not unexpected when we switch from the uncensored to the censored adoption sample. The bias is bigger than in our previous samples and varies between 11 and 21 percent. Probably because of the smaller samples, the censoring bias is rather imprecisely estimated, and never statistically significant (see columns (7), (8) and (9)).

In the remaining panels we evaluate the various solutions to the censoring bias. Compared to the uncorrected regression results using the censored adoption sample the three approaches produce (almost always) higher mobility estimates and thus appear to remove the downward bias. Compared to the results in Table 2.2 the estimated impacts of parental schooling drop and inherited abilities and assortative mating seem to play a more important role for mothers than for fathers. There is one exception. In the last panel where we eliminate all school-going aged children, we obtain a coefficient on mother's schooling that is statistically significant and almost of the same size as the coefficient on father's schooling. In columns (7), (8) and (9) we report on the differences between the estimates using the three approaches and the estimates without censoring. These differences are

larger than those differences reported in Table 2.2 using the sample of own birth children. But since resulting differences are all statistically insignificant, it is difficult to draw firm conclusions about the validity of each solution.

2.5 Can we treat expectations as realizations?

Our results in Table 2.2 suggest that parental expectations fix the censoring problem quite well.⁹ This is by no means a trivial result. In a recent paper Antonovics and Goldberger (2005 p.1739) express their doubt regarding this particular correction method. We therefore perform additional robustness checks to see how sensitive the parental expectations solution is to a number of potential threats: the number of censored observations, sample selection and prediction quality.

The degree of censoring

Our first concern is that the expectation method might work well because the number of censored observations in our sample is relatively small. In Section 2.2 we showed that the bias introduced by replacing censored observations with parental expectations depends on the correlation of parents' schooling with parental prediction errors (η_t) and the degree of censoring (d_t); that is,

$$\tilde{\beta}_{LS} - \beta = \text{cov}(d_t \eta_t, S_{t-1}) / \text{var}(S_{t-1}).$$

This is an expression we can actually test: least squares estimation of the regression of $d_t \eta_t$ on parental schooling. To see whether our results are sensitive to the number of censored observations, we estimate the bias of the replacement method on samples where we gradually increase the number of censored observations. We do this by calculating how

⁹In this Section we restrict our attention to the own birth results. Since the bias estimates in our adoption analysis are too imprecise and not informative about the preferred correction approach, we have decided to ignore the adoption results, at least, when we evaluate the three correction methods.

many children would still be in school if we had observed them some years before 1992. For example, if a mother, who reports in 1992 that her child, born in 1967, completed 15 years of schooling, were interviewed in 1984 we recode the same child as being censored, assuming he/she left school in 1988 ($1967+6+15$). In 1984 the same mother would have reported that her child had 11 years of schooling, assuming that children start school at age 6 and have uninterrupted school careers.

The first panel of Table 2.4 contains the estimates of the bias when using the replacement method, for increasing numbers of censored observations, with additional controls for age and gender of the child. Up to censoring percentages of 60, we find that all the bias estimates are statistically insignificant and virtually zero, confirming our baseline result that the replacement method yields consistent mobility estimates. Up to censoring percentages of 90, the bias is negative but small, and often statistically insignificant. The procedure to replace the censored observations with expectations is statistically rejected, but only at the margin. Only when the percentage of censored observations becomes very large, the corresponding method to adjust for censoring fails. The slopes are negative and statistically significant. Would we fully rely on parental predictions, the implication is that the corresponding intergenerational mobility estimates are biased downwards. The negative bias further suggests that expectations regress to the mean faster than realizations do.

In the second and third panel we also show results for the censored regression and elimination models. In case of the censored regression approach, we find for small censoring percentages that the estimated bias is somewhat larger than the bias reported in the previous panel. When we increase censoring percentages, we find that the bias consistently falls. For censoring percentages around 50 percent the bias goes to zero and then becomes negative for samples where the majority of the children is still in school.¹⁰

¹⁰This pattern is consistent with a bimodal schooling distribution. Arabmazar and Schmidt (1982) investigate the inconsistency of the related Tobit estimator as a consequence of different non-normal distributions. They find, that the bias due to non-normality depends on the degree of censoring. They do, however, not investigate the consequences of a bimodal distribution. If we assume a bimodal distribution of years of schooling, our simulation results –not reported in the chapter– bear out

Table 2.4: Estimating the bias for increasing censoring percentages

	Samples with increasing numbers of censored observations							
	20-30%	30-40%	40-50%	50-60%	60-70%	70-80%	80-90%	100%
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
CENSORED OBSERVATIONS REPLACED WITH PARENTAL EXPECTATIONS								
Mother's schooling	-0.01 (0.01)	-0.02 (0.02)	-0.02 (0.02)	-0.02 (0.02)	-0.02 (0.02)	-0.03 (0.02)	-0.04 (0.02)	-0.05** (0.02)
Father's schooling	0.00 (0.01)	-0.01 (0.01)	-0.02 (0.01)	-0.02 (0.01)	-0.02* (0.01)	-0.03* (0.01)	-0.03* (0.01)	-0.04*** (0.02)
<i>N</i>	4097	4097	4097	4097	4097	4097	4097	4097
<i>N^c</i>	874 ^a	1285	1712	2280	2859	3157	3388	4097
CENSORED REGRESSION MODEL								
Mother's schooling	0.05*** (0.01)	0.06*** (0.02)	0.06*** (0.02)	0.04 (0.03)	0.04 (0.03)	-0.01 (0.04)	-0.06 (0.04)	x
Father's schooling	0.03*** (0.01)	0.02** (0.01)	0.01 (0.01)	0.00 (0.01)	-0.02 (0.02)	-0.06*** (0.02)	-0.09*** (0.02)	x
<i>N</i>	4097	4097	4097	4097	4097	4074	4006	x
<i>N^c</i>	874	1285	1712	2280	2859	3134	3297	x
EXCLUDING ALL CHILDREN YOUNGER THAN 25								
Mother's schooling	0.04** (0.02)	0.07** (0.04)	0.03 (0.06)	0.02 (0.18)	x	x	x	x
Father's schooling	0.01 (0.01)	0.02 (0.02)	-0.01 (0.03)	-0.00 (0.06)	x	x	x	x
<i>N</i>	2990	1558	852	240	x	x	x	x

All regressions include controls for the child's age and gender. Standard errors are in parentheses. * significant at 10% level, ** significant at 5% level, *** significant at 1% level.^aNumber corresponds to true number of censored observations in 1992.

that the inconsistency of the maximum likelihood estimator is positive when about 25 percent of the observations is (right) censored, and negative when about 75 percent of the observations is censored.

In case we exclude all children below 25 from our sample, we find for small percentages that the bias is positive, statistically significant and comparable to the bias of the censored regression model. Together with falling sample sizes, the bias declines for censoring percentages up to 60 percent. For samples where more than 60 percent of the observations is censored, all children are below the age of 25 and the elimination method no longer works. Overall, we believe that of the three different solutions to the censoring problem, the replacement method appears to be the least sensitive to the number of censored observations.

Prediction quality

Our second concern is that the replacement trick might work because of the nonrepresentative nature of the WLS. The WLS only collects information on high school graduates and (because of that) systematically under-samples the lower educated individuals. If more schooled parents form more accurate expectations about their children's schooling, it is possible that our observation –the best approach is to replace censored observations with parental expectations– is driven by the sample design of the WLS, and does not hold in other data sets.

Other data sources with panel information on completed schooling and expected schooling (when the same children were still in school) would resolve part of this external validity discussion. These two measures, however, are rarely collected in large systematic data sets. It is still possible to get an idea whether the mechanism of more schooled parents forming more accurate expectations is present among WLS parents. We first ask ourselves whether WLS parents can perfectly predict their child's education. Figure 2.1 shows a histogram of the difference between parental expectations and realizations.

Although for more than 35 percent of the children parental expectations coincide with realizations, there is quite some variation in how well parents can predict their child's schooling. To explore this variation further, Figure 2.2 shows scatter diagrams and lowess

estimation of the relation between the prediction error and mother's and father's schooling. This figure indicates that there is no strong relation between parental schooling and the difference between expected and completed schooling of the child. We also estimate how the absolute prediction error depends on parental schooling, age and gender of the child. We observe that the coefficient (with corresponding absolute t ratio between brackets) is 0.013 [0.76] for mother's schooling and -0.021 [1.88] for father's schooling. If we look at mothers, there appears to be no evidence that more schooled mothers make better predictions. If we look at fathers, we do find that prediction quality improves with years of schooling, but it is only at the margin. These weak correlations, we think, raise the generalizability of our findings.

Figure 2.1: Difference between parental expectations and completed years of schooling

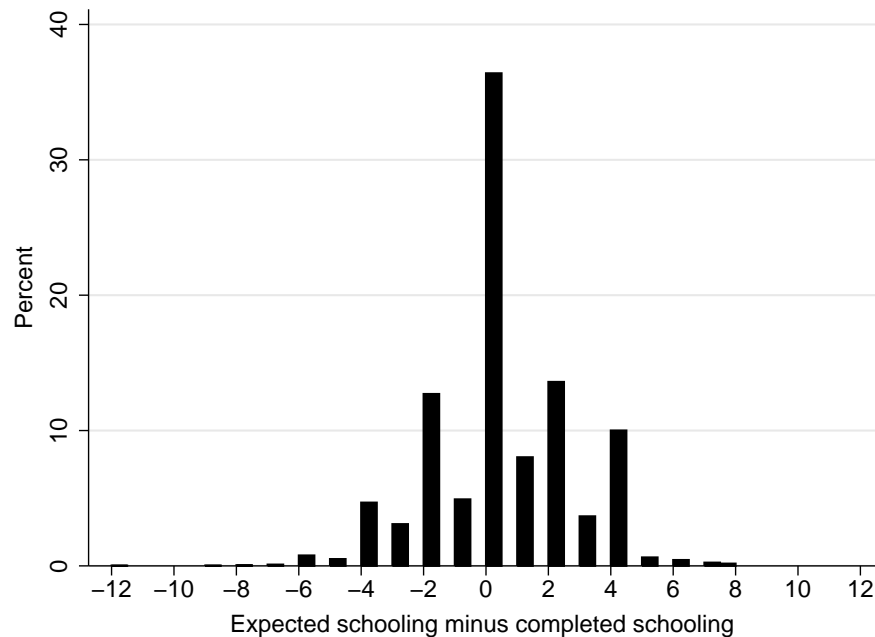
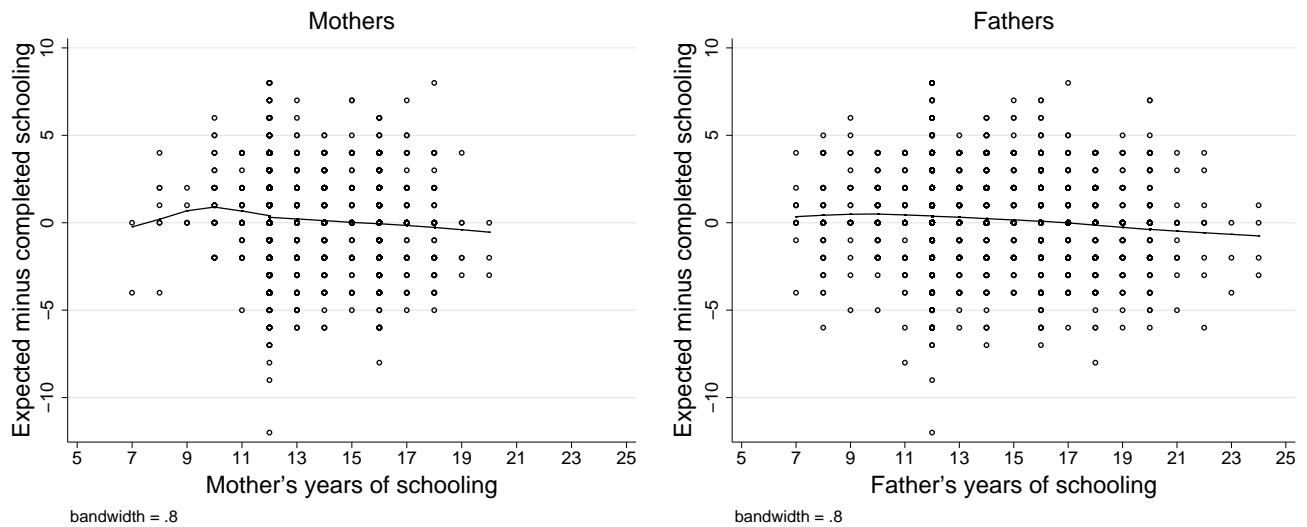


Figure 2.2: Lowess smooting: relation between parental schooling and prediction errors



Sample of selected children

So far, our findings with respect to the maximum likelihood and elimination approaches are based on the full sample of children while the results using the replacement method are generated using the smaller sample of one randomly selected child per family. It might be that the results reported in the previous sections are sensitive to the specific sample that is used. To check whether this is the case, Table 2.5 shows the same analyses as in Table 2.2 but now all results are based on the sample of selected children.

As can be seen in Table 2.5 the results are not very sensitive to the specific sample that is used to generate the results. Ignoring the censored observations give estimates that are downwardly biased. The maximum likelihood approach overestimates the effect of mother's and father's schooling, all the estimated biases are significantly different from zero. Also the elimination approach produces positively biased estimates, only for father's the differences are not significantly different from zero. The results whereby we replace the censored observations with parental expectations are the same as in Table 2.2 and show no significant bias.

Table 2.5: Estimates of the effect of parents' schooling on own birth children's schooling—sample of selected children

	Mobility Estimates without Censoring (WLS 2004)			Mobility Estimates with Censoring (WLS 1992)			Estimated Differences		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mother's schooling	0.45*** (0.02)		0.24*** (0.02)	0.42*** (0.02)		0.22*** (0.02)	-0.03*** (0.01)		-0.02 (0.01)
Father's schooling		0.35*** (0.01)	0.27*** (0.01)		0.32*** (0.01)	0.25*** (0.01)		-0.03*** (0.01)	-0.02*** (0.01)
<i>N</i>		4,097			4,097				
<i>N^c</i>		0			874				
CENSORED REGRESSION MODEL									
Mother's schooling				0.53*** (0.02)		0.30*** (0.03)	0.08*** (0.01)		0.05*** (0.01)
Father's schooling					0.39*** (0.01)	0.30*** (0.02)		0.04*** (0.01)	0.03*** (0.01)
<i>N</i>					4,097				
<i>N^c</i>					874				
EXCLUDING ALL CHILDREN YOUNGER THAN 25									
Mother's schooling				0.50*** (0.03)		0.28*** (0.03)	0.05*** (0.02)		0.04*** (0.02)
Father's schooling					0.36*** (0.02)	0.28*** (0.02)		0.01 (0.01)	0.01 (0.01)
<i>N</i>					2,990				
<i>N^c</i>					167				
CENSORED OBSERVATIONS REPLACED WITH PARENTAL EXPECTATIONS									
Mother's schooling				0.45*** (0.02)		0.23*** (0.02)	-0.00 (0.01)		-0.01 (0.01)
Father's schooling					0.35*** (0.01)	0.27*** (0.01)		0.00 (0.01)	0.00 (0.01)
<i>N</i>					4,097				
<i>N^c</i>					874				

All regressions include additional controls for the child's age and gender. Standard errors (in parentheses) allow for correlation within families. * significant at 10% level, ** significant at 5% level, *** significant at 1% level.

Patience versus impatience

Since the replacement method is not as disconcerting as Antonovics and Goldberger say it is, it is interesting to see what happens if a researcher is very impatient and wants to estimate mobility when none of the children has finished their schooling. Table 2.6 compares the mobility estimates on the uncensored sample with the results obtained when a researcher would have used the 1975 sample and replaced all observations by parental expectations.¹¹ For the sample of own birth children we find mobility estimates of 0.19 and 0.23 for mothers and fathers, respectively. For adoptees the effect of mothers schooling drops to 0.11 and is no longer significantly different from zero. The effect of father's schooling does not change much and remains statistically significant. Compared to the uncensored mobility results, these estimates are statistically but not substantially different, which is quite remarkable given that schooling expectations were measured when almost all children were still in primary school.

Maximum likelihood and the elimination approach

The sensitivity analysis thus far has concentrated on mechanisms that could possibly invalidate the method to replace censored school observations with parental expectations. But, perhaps it is also informative to understand why the maximum likelihood and elimination approach produce biased mobility estimates, even though the bias as reported in Table 2.2 is not substantial.

We begin with the censored regression model. One likely candidate to explain the upward bias of the maximum likelihood approach would be a normality violation. It is unlikely that schooling is normally distributed –the more appropriate distribution of the child's completed education is bimodal with peaks around 12 and 16 years (see also footnote 10).

¹¹The other two methods do not work with samples where all the observations are censored.

Table 2.6: Patience versus impatience

	Estimates without Censoring (WLS 2004)		Estimates with Expectations (WLS 1975)	
	(1)	(2)	(3)	(4)
Mother's schooling	0.24*** (0.02)	0.08 (0.06)	0.19*** (0.02)	0.11 (0.16)
Father's schooling	0.26*** (0.01)	0.19*** (0.04)	0.23 *** (0.01)	0.26** (0.10)
<i>N</i>	14,524	520	4,097	52
SAMPLES				
Own birth children	×	–	×	–
Adopted children	–	×	–	×

All regressions include additional controls for the child's age and gender. Standard errors are in parentheses; * significant at 10% level, ** significant at 5% level, *** significant at 1% level.

Another possibility is that heteroskedasticity is causing the inconsistent estimates. Using the uncensored 2004 sample we can test whether the normality and or homoskedasticity assumptions are violated. The results show that the null hypothesis of homoskedasticity is not rejected but that the normality assumption is indeed rejected.¹²

A candidate to explain the inconsistencies caused by the elimination procedure would relate to the fact that by eliminating all children below 25 we are left with a sample consisting of parents who chose to have children at a relatively early age. These parents are likely different in both observed and unobserved characteristics which can cause the up-

¹²The tests for normality and homoskedasticity are performed on the specification including both mother's and father's schooling as regressors. The p-value of the Breusch-Pagan/Cook-Weisberg test is equal to 0.366, the null hypothesis of homoskedasticity is therefore not rejected. The p-value of the skewness/kurtosis tests for normality is equal to 0.000, the null hypothesis of normality is therefore rejected

ward bias we observe in Table 2.2. Parents who choose to have children at an early age, for example, are mostly lower educated. If mobility is lower at the lower end of the distribution (Oreopoulos *et al.* 2006) the elimination of mostly children from higher educated parents would lead to an estimate of the mobility parameter that is too high.

2.6 Conclusion

Recent mobility studies that make a distinction between causation and selection often rely on samples in which information on the child's completed schooling is not always available. Unfortunately, solutions offered to handle censored samples do not always work, and should be further scrutinized.

This is what we do in this chapter. We first estimate the impact of mother's and father's schooling on child's schooling using censored and uncensored samples of own birth children and adoptees, and then investigate the consequences of three different methods that deal with censored observations: maximum likelihood approach, replacement of observed with expected years of schooling and elimination of all school-aged children.

Our basic result is that, net of assortative mating effects, positive parental schooling effects fall only little for fathers, but much more for mothers when we move from uncensored samples of own birth children to adoptees. This result appears to be fairly robust to the introduction of censored observations and the application of three correction methods. Parental schooling effects fall, but not by much, when mobility models are estimated on censored samples and rise, again not by much, when censored observations are tackled by either three correction methods. Of the three methods, the one that treats parental expectations as if they were realizations performs best. This result depends, however, on the degree of censoring. For samples that are largely incomplete the method does give a small (negative) bias.

Our results suggest that it doesn't matter (much) whether researchers are patient or

impatient: whether we fully rely on parental expectations, or whether we use realizations measured 30 years later, the mobility estimates are not substantially different. However, to draw more general conclusions from these remarkable findings we think that more research is needed on how parents form expectations and on how the replacement method works using different identification strategies on other data sets.

Chapter 3

Birth order, family size and educational attainment

3.1 Introduction

Many studies indicate that birth order and family size are important determinants of educational outcomes of children. Family size and birth order are strongly related, although family size differs between children from different families, while birth order differs between children within a family. In previous research often no clear distinction between sibship size and birth order is made and estimated effects of sibship size could be picking up the effect of birth order and visa versa. Because of the strong relation between birth order and family size this chapter estimates the effects of both family background components on years of schooling, using the Wisconsin Longitudinal Study .

The relationship between family size and educational attainment can be the result of constraints on parental resources. When there are capital market imperfections and parents have many children, they can, for a given income, invest less in each child than if they have fewer children. This can cause a negative relationship between family size and educational attainment (Becker (1981)). Also numerous empirical studies have found a negative relationship between the number of siblings, and future economic and educa-

tional achievements.

This negative relationship between family size and educational achievements is however not necessarily proof of a negative effect of the number of children. The number of children is a choice variable of the parents and it might be that certain characteristics of parents, such as their educational attainments, affect both the number of children as well as the educational attainments of those children.

This chapter uses an instrumental variable (IV) approach to identify the effect of sibship size on years of schooling of a first-born child. Two sources of exogenous variation in the number of children are used; twins at last birth, and the preference of parents for a mixed sibling sex composition. Like many studies, this chapter finds a negative correlation between the number of children and child's years of schooling. This negative correlation declines however when control variables like parental schooling and birth order dummies are added. This signals that the observed negative correlation might not be causal. The IV results indeed are no longer significantly negative, but positive and insignificant. Although the standard errors are not small, these results indicate that exogenous variation in the number of children does not have a significant negative effect on the educational attainment of a first-born child.

Since including birth order dummies makes the coefficient on family size insignificant, it might be that the negative relation between family size and years of schooling was just picking up the impact of birth order. Models from psychology predict a decline in intellectual environment with birth order, which can cause a negative effect of birth order on educational achievements (Zajonc (1976)). Economists emphasize the constraints on available parental time and resources, which can cause children's schooling outcomes to decline with birth order (Becker (1981), Behrman (1997)).

This chapter also identifies the effect of birth order, by estimating the effect on years of schooling separately for families with two, three, four or five children. Family size is correlated with birth order, and not taking this into account would give estimates of

the birth order effects which might confound birth order with family size, or with family characteristics that are correlated with family size. Estimating the effect of birth order separately for families with a different number of children avoids this problem. For all family sizes examined in this chapter, birth order turns out to have a significant negative effect on educational attainment. This decline in years of schooling with birth order turns out to be approximately linear.

Recent studies have also looked at the impact of family size (Cáceres-Delpiano (2006), Angrist *et al.* (2007)), birth order (Kantarevic and Mechoulan (2006)) or both (Black *et al.* (2005B), Conley and Glauber (2006)). The findings in this chapter are very similar to the findings of most of these papers in the sense that the results also show no significant causal impact of family size on years of schooling but a significant decline in educational attainment with birth order. The main contribution of this chapter is that it goes beyond estimating the impact of birth order and family size, by investigating two potential mechanisms behind the estimated effects, competition between closely spaced siblings and the allocation of parental resources.

Different theories predict that birth order effects will be smaller when the time between births is larger. This chapter investigates the effect of child-spacing, taking into account the possible endogeneity of the space between births, first by including family fixed effects and secondly by using the presence of twins as instrument.

To investigate whether restrictions on parental resources are behind the birth order effects, this chapter exploits information about the amount of money parents report to have given to their children. This information about financial transfers to children makes it possible to investigate whether there is a causal impact of the number of children on the probability that a child receives money from his parents and whether this probability and the amount received vary with the birth order of the child.

The results show that the negative effect of birth order does not vary significantly with the average space between births. This results is robust to including family fixed effects

and the use of the presence of twins as instrument. This chapter further finds that earlier born children have a higher probability of receiving money from their parents and they receive a larger amount. These findings indicate that the allocation of parental resources is a potential mechanism behind the birth order patterns observed in children's schooling outcomes.

The plan of the chapter is as follows. Section 3.2 will give an overview of the theoretical and empirical literature. Section 3.3 continues with a description of the data used. Section 3.4 gives the results on respectively the impact of family size and birth order. Section 3.5 investigates potential mechanisms behind the birth order effects and finally Section 3.6 concludes.

3.2 Theoretical and empirical background

Family size and educational attainment

There is an extensive theoretical literature about the trade-off between child quality and quantity, dating back to the models of Becker and Lewis (1973) and Becker and Tomes (1976). The idea behind these theoretical models is that if parents have more children, investing a certain amount in per-child quality, for example their education, is more expensive, than if they have fewer children. When there is an (exogenous) increase in the number of children, the total cost of investing a certain amount per child becomes higher and for a given budget constraint parents will lower the investment in per-child quality. This indicates that there is a negative relation between child quantity and child quality.

However, parents not only have an influence on child quality through investment of resources, but also through transmission of their endowments. If parents with lower endowments have a higher preference for child quantity than parents with higher endowments, and therefore also have more children, this can cause a negative correlation between child quantity and child quality, by way of the effect of parental endowments on child quality.

Children of parents with low endowments will in this case have on average more siblings and a lower educational attainment, even though there may be no causal effect of the number of children on educational attainment.

The empirical studies investigating the relation between child quantity and quality usually perform a least squares regression of economic and educational outcomes on sibship size and other socioeconomic background variables. Most of these studies have found a significant negative relation between the number of children and the educational achievements of those children. Examples of these studies are Belmont and Marolla (1973), Blake (1981) and Hanushek (1992).

These studies do however not take the possible endogeneity of the number children into account. Recently the effect of family size on educational achievement is investigated using an instrumental variable approach. Black *et al.* (2005B) use multiple births as instruments for the number of children, to investigate the effect of sibship size on children's education in Norway. They find a negative correlation between family size and educational attainment, but when they include control variables such as birth order dummies, and when they use the twin births as instruments, they find no significant negative effect of the number of children on educational attainment. They look only at Norwegian data and their results might not generalize to other countries. The study in this chapter applies a similar methodology by using twins as instrument to identify the effect of family size using data from the United States. It further complements their paper by using the sex mix of the first two children as instrument for family size, next to the twins instrument¹. By using both instruments the effect of family size can be identified using different sources of exogenous variation.

Conley and Glauber (2006) and Cáceres-Delpiano (2006) both use U.S. census data to estimate the effect of the number of children. Conley and Glauber use exogenous

¹Black *et al.* (2005B) briefly mention some results using the sex mix of the first two children as instrument. They find a significant positive effect of the number of children on years of education, and they do not find the magnitude of their estimate credible.

variation in sibling sex composition to estimate the causal effect of family size. They find no effect of family size for first-born boys but they do find effects for second born boys; an increase in the number of siblings reduces their likelihood of private school attendance and increases grade retention.

Cáceres-Delpiano uses multiple births as an instrument for family size. He finds significant effects on investment in child quality in the sense that a larger family reduces the likelihood that the oldest child attends private school, reduces the mother's labor force participation, and increases the likelihood that parents divorce. He does not find a significant impact on child outcomes, measured by grade retention.

These two studies do not investigate the effect of the number of siblings on the final schooling outcomes, and they only look at young children who still reside with their parents. Both studies restrict their samples to families where the total number of children living in the household is equal to the number of children the mother has ever given birth to. However, since the mothers are quite young this reported family size need not be equal to the desired or the eventually completed family size. In contrast the data set used in this chapter contains completed family size and completed schooling outcomes for all children in a family, whether they are still residing with their parents or not.

Angrist *et al.* (2007) use exogenous variation in family size both due to twin births and due to the preference for a mixed sibling sex composition. They also include interactions of these instruments with ethnicity, exploiting the difference in the effects of the instruments for different ethnic groups. Using Israeli Census data they find negative and significant OLS coefficients, but using the instruments they find no significant effect of the number of children on educational and labor market outcomes.

The studies by Angrist *et al.*, Black *et al.* and Cáceres-Delpiano use twins at second or at a higher birth as instrument for the number of children. Families who had, for example, twins at second birth and who have more than three children in total would however likely have had more than three children anyway, even if they would not have had twins at

second birth. The investigation in this chapter uses instead twins at last birth and this prevents using multiple births as exogenous variation in the number of children, while the parents wanted more children anyway.

Birth order and educational attainment

Different theories predict a significant effect of birth order on future educational and economic achievements. An important model in the psychological literature is the confluence model, which is discussed by Zajonc (1976). The confluence model predicts the intellectual growth of a child to depend on its intellectual environment, whereby the intellectual environment is defined as a function of the average of the absolute intellectual levels of its family members. A first child enters an environment with two adult parents, so the average intellectual level is high. A second born enters a lower intellectual environment than the first-born, because his older sibling is part of the intellectual environment, and decreases the average intellectual level. A third child enters an even lower intellectual environment than the second born because of two older siblings etcetera. Since the intellectual environment declines with birth order, this model predicts a negative relation between birth order and educational attainment. Zajonc however stresses the effect of child-spacing, the larger the age gap between two siblings the smaller is the difference between the two children in their intellectual environment, and the smaller is the effect of birth order.

Economic theories underline the restrictions on available parental time and resources (Becker (1981), Behrman (1997)). Later born children probably spend on average less time with their parents than earlier born children. A first-born does not have to share his parent's time with siblings until a second child is born. The second child has to share parental time with its older sibling, but only from the moment a third child is born, he or she has to share with another sibling. So in particular if time spent in the early childhood is more important than time spent with children when they are older, later born children are disadvantaged compared to earlier born children, because parents cannot spend as much

time with later born children, as they did with earlier born children.

Not only restrictions on the amount of available time can cause a negative effect of birth order, but also limitations on the amount of resources available to invest in the human capital of the children may cause a negative effect. If parents simultaneously decide on the number of children and the quality of those children, that is how much to invest in each of their children, no birth order effect would be expected, unless parents explicitly favor earlier born children over later born children. In the quantity-quality model it is implicitly assumed that parents take the cost of investment in their children into account when deciding on the number of children. However, if parents are not enough forward looking, or when it is difficult for parents to estimate the actual amount of resources necessary to invest in their children's education, later born children may find that most resources have been depleted by older siblings, at the moment they need resources to finance their education.²

Empirical studies have found different effects of birth order. Belmont and Marolla (1973) examine, next to the effect of family size, the effect of birth order for given family sizes. They find a steady decline in the average score on a military examination with birth order. Blake (1981) also examines the effect of birth order and does not find a systematic difference in educational attainment between first and last borns, and middle born children.

Hauser and Sewel (1985) use the sample of 9000 Wisconsin high school graduates of 1957 and their siblings. In this sample also no significant or systematic effects of birth order on educational attainment are found. A drawback of this sample is that all graduates have at least high school. Later in this chapter the offspring of these graduates will be examined, and the variation in years of education is larger among the children in this sample, there are both children with more than a high school degree as children with less than a high school degree.

²A binding budget constraint can, however, also favor later born children over earlier born children. If parents have an upward sloping earnings profile and there are capital market imperfections, parents can even have more resources available to invest in later born children, than they had available for their earlier born children.

Hanushek (1992) studies the effect of family size and birth order on children's scholastic performance. He finds that earlier born children perform better, but that this is entirely because earlier born children have a higher probability of being in a small family. For fixed family sizes he finds no significant differences between first and last born children.

Behrman and Taubman (1986) study the birth order effect for the adult offspring of the twins in the National Academy of Science/National Research Council (NAS/NRC) sample. They find in contrast to the previous three studies, significant positive effects of being an early born on (age-adjusted) schooling. Behrman and Taubman include the number of children next to other family background characteristics in their specification, because the estimates might confound birth order with family size. This does however not solve the problem entirely, it is still difficult to separate the effect of birth order and family size, and the number of children is likely endogenous. The study in this chapter investigates the effect of birth order for families with two, three, four and five children separately. In this way estimated effects of birth order cannot confound with the effect of family size.

Black *et al.* (2005B) use data from Norway. Next to the investigation of the effect of the number of children on years of schooling, they also examine the effect of birth order for different family sizes separately. By including birth order dummies they find a significant negative effect of birth order on educational attainment, for all family sizes examined.

Kantarevic and Mechoulan (2006) use the Panel Study of Income Dynamics and find that first-born children have a significantly higher educational attainment. They also estimate birth order effects by family size but due to small sample sizes the estimates become imprecise for larger families. This chapter will also investigate the effect of birth order on educational attainment for children from the United States. The sample size used in this chapter is larger than the sample used by Kantarevic and Mechoulan, which gives the possibility to estimate all specifications by family size and to get more precisely estimated coefficients. It further adds to the previous two papers by investigating some of the

mechanisms which could be behind the estimated birth order effects.

3.3 Data

The data set used in the empirical study in this chapter comes from the Wisconsin Longitudinal Study (WLS), which is a long-term study of a random sample of 10,317 men and women who graduated from Wisconsin high schools in 1957. Survey data were collected from the high school graduates in 1957, 1964, 1975, 1992, and 2004 and from a randomly selected sibling of the graduate in 1977, 1994, and 2005. The data contains, among others, information about social background, schooling, intergenerational transfers and relationships, and family characteristics. This chapter will mainly use information from the last two rounds of the questionnaire since these contain updated information about completed years of schooling of all of the children of the graduates and of all the children of the selected siblings. This is important since the analysis will look at the effect of birth order and family size on the completed schooling outcomes of these children.

For the study in this chapter the data set is reduced to include only married respondents, with a minimum of two and a maximum of five children. This selection of the data set is necessary to be certain that children lived together during their childhood. When parents get divorced it happens that some children live with their mother and the other children live with their father, and in these kinds of circumstances the number of children reported in the survey is not equal to the number of children that lived with the respondent during the childhood of the children.

Including only married respondents reduces the number of respondents to 6,980 graduates and 4,263 siblings.³ Eliminating all respondents with less than two children and more than five children further reduces the data set to 5,762 graduates and 3,341 selected

³All regressions in this chapter have also been performed without this selection of the data set and the results were very similar to those reported in the chapter. To avoid complications in interpreting the results, for example when children did not live together during their childhood because their parents did not live together, only results for the selected sample are presented.

siblings. Families with more than five children are excluded, because in the birth order analysis families of different sizes are examined separately, and for families with more than five children this is not possible due to sample size problems. The sample is further restricted to families with only biological children, because it might be the case that adopted children are treated differently by their adoptive parents than biological children. Moreover, little is known about the age of adoption and since the effect of birth order is examined in this chapter, adopted children who become part of a family some years after they were born might complicate the analysis.

In 2004/2005 4,877 of the graduates and 2,562 of the selected siblings responded to the follow up questionnaire. The final sample of children for whom we have completed schooling outcomes consists of 17,113 children from 5,990 families.⁴ These children were born between 1943 and 1981, and 85% of these children was born between 1960 and 1975. On average they have completed 14.58 years of schooling. Table 3.1 gives some descriptive statistics of the sample.⁵

3.4 Results

Family size and educational attainment

In previous studies the relation between the number of children and the educational attainment of a child is usually estimated using an ordinary least squares specification. In this section the possible endogeneity of the number of children will also initially be ignored, and the following equation will be estimated three times, each time with a different set of control variables.

$$E_{child} = \alpha \cdot children + X \cdot \beta + \varepsilon \quad (3.1)$$

⁴There are some children below the age of 23 in 2004 and these children might still be in school. In the analysis we eliminate these observations. In the sample in this chapter less than 2% is below the age of 23. It is unlikely that eliminating these observations can cause a significant bias in the estimates.

⁵Since the sample contains multiple children from one family, these observations are not independent. All the specifications are therefore estimated with clustered standard errors.

First X only contains dummies for the age and gender of the child, then the years of schooling of the father and the mother are included, and finally a set of dummies indicating whether the child is a second, third, fourth or fifth born are added to the specification.

Table 3.1: Descriptive statistics

	Mean	std. dev.
Child's years of schooling	14.58	2.31
Age child in 2004	38.12	5.48
Gender (girl=1)	0.49	0.50
Number of children	3.33	0.99
Father's years of schooling	13.57	2.66
Mother's years of schooling	12.94	1.78
Twins at last birth	0.01	0.11
Twin present	0.03	0.17
First two children of same sex	0.51	0.50
Average space (in months)	34.03	16.40
Child received money from parents ^a	0.46	0.50
Amount child received from parents ^b	7200.81	15113.82
Number of families	5,990	
Number of observations	17,113	

^{a,b}Based on a sample of children from families in which all children were older than 25 in 1992. ^aN=5695 ^bN=4059

The highest level of education, measured in years of schooling, is used as dependent variable and the number of children ranges from two to five. Table 3.2 shows the estimation results of this equation.

As can be seen in column 1 in Table 3.2, the coefficient on the number of children is significantly negative. The magnitude of the coefficient declines however when paternal and maternal education are added as control variables. This signals that the number of children is endogenous, and that the school choice of parents is somehow related to their decision about the number of children. Another possibility is that the negative effect of family size is picking up the negative effect of birth order. Therefore in column 3 birth

order dummies are added, and indeed the coefficient on the number of children declines even further and is no longer significantly different from zero. The birth order dummies are however all significantly negative. It is not possible though to conclude from this table what the causal effect of the number of children is. It is also difficult to separate the effect of birth order and sibship size in column 3, because the number of children and the birth order dummies are strongly correlated.⁶

Table 3.2: The effect of the number of children on child's years of schooling

	(1)	(2)	(3)
Number of children	-0.24*** (0.02)	-0.11*** (0.02)	-0.03 (0.02)
Fathers years of schooling		0.27*** (0.01)	0.26*** (0.01)
Mother's years of schooling		0.23*** (0.01)	0.23*** (0.01)
Second born			-0.26*** (0.03)
Third born			-0.39*** (0.05)
Fourth born			-0.50*** (0.06)
Fifth born			-0.49*** (0.11)
Number of families	5,990	5,990	5,990
Number of observations	17,113	17,113	17,113

All regressions include dummies for age and gender of the child.

Standard error (in parentheses) allow for correlation within families.

*significant at 10% level, **significant at 5% level, ***significant at 1% level.

To identify the causal impact of the number of children on years of schooling of the first-born child, twins at last birth and the sex mix of the first two children will be used

⁶The correlation between a continuous variable birth order and the number of children is equal to 0.46. The continuous birth order variable has the value zero for a first-born child, the value one for a second-born child etc. The variable ranges from zero to four.

as instruments.⁷ The variable twins at last birth takes on the value one in three situations; when there are three children in the family and the second birth is a twin, when there are four children and the third birth is a twin and the final situation is when there are five children and the fourth birth is a twin.

For twins at last birth to be a valid instrument for the number of children, it should have no separate effect on the educational attainment of the first-born child. If child-spacing has an effect on educational attainment, using twins at last birth as instrument could be a potential problem, because in the case of twins the space is zero. If the age gap between two children is smaller the competition between siblings for time and parental funds is likely higher, which can have a negative effect on the education of a child, separate from the effect via the number of children. Estimating, for fixed family sizes, the effects of the number of months between two subsequent births on years of schooling of the first-born child, gives coefficients which are all not significantly different from zero.⁸ This indicates that the space between births has no significant effect on educational attainment. Later in this chapter the effect of child-spacing will be investigated more thoroughly.

Another potential problem is that the probability of having twins at last birth might be affected by certain characteristics of parents, such as their educational attainment. Results from linear probability models show however that the education of both the father and the mother have no significant effect on the probability of having twins at last birth.⁹ These results indicate that the instrument twins at last birth has no separate effect on years of schooling of the first-born child, although this remains of course an untestable assumption.

The second instrument takes on the value one when the first two children are two boys or two girls. If the first two children are of the same sex, parents are significantly

⁷See Rosenzweig and Wolpin (2000) for a discussion on the use of these instruments.

⁸Regressions are by family size and include the number of months between the second and third child (space23), between third and the fourth child (space34) and between fourth and fifth child (space45). All regressions include dummies for age and gender of the child and controls for father's and mother's years of schooling. Results for three child families: space23: 0.002(0.002), four child families: space23: -0.002 (0.003), space34: 0.001 (0.002), and for five child families: space23: -0.004 (0.007), space34: -0.005 (0.006), space45: -0.005 (0.004).

⁹The coefficient on father's education is equal to -0.0008 (0.0006) and the coefficient on mother's education is equal to 0.0007 (0.0007).

more likely to have another child, because of the widely observed preference of parents for a mixed sibling sex composition (Ben-Porath and Welch (1976), Angrist and Evans (1998)).¹⁰ For the sex mix of the first two children to be a valid instrument, it should also have no effect on the educational attainment of a child, separate from the effect via the number of children. Some papers indicate that the number of brothers and sisters can have an effect on educational outcomes. Butcher and Case (1994) find that, while controlling for family size, women raised with only brothers have a significantly higher educational attainment than women with any sisters. Hauser and Kuo (1998) report however that no effect of the sibling sex composition on educational attainment for the sample of 9000 Wisconsin high school graduates has been found. They go on to perform a similar analysis as Butcher and Case on three survey data sets, and they do not find any evidence for an effect of the gender composition of sibships on years of completed schooling. Also Kaestner (1997) does not find significant effects of sibling sex composition on the educational attainment of white males and females.

For the data set used in this chapter, two dummy variables are added to the specification in equation (3.1), one for having only sisters and one for having only brothers, children with both brothers and sisters are the control group. Both dummy variables have an insignificant effect on the educational attainment of the first-born child.¹¹ These results indicate that the gender composition of sibships does not have a significant effect on educational attainment. So in the subsequent analysis it is assumed that the sex mix of the first two children has no effect on the schooling of the oldest child, separate from the effect via the number of children. The following model will be estimated

$$E_{child} = \alpha_0 + \alpha_1 \cdot children + X \cdot \beta + v \quad (3.2)$$

$$children = \delta_0 + \delta_1 \cdot Z + X \cdot \psi + \eta \quad (3.3)$$

¹⁰Angrist and Evans also use the sex mix of the first two children as an instrument for the number of children, to test the effect of childbearing on labour supply.

¹¹The coefficient on 'only sisters' is equal to -0.11 (0.08) and the coefficient on 'only brothers' is equal to 0.13 (0.08).

Z is the instrumental variable for the number of children and E_{child} is the completed years of schooling of the first-born child. Only the effect of the number of children on years of schooling of the oldest child is examined, because children who are born as second, third, fourth or fifth and who are part of a twin can be affected directly by the instrument twins at last birth. Children who are part of a twin have, for example, often a lower birth weight than children not part of a multiple birth and this might affect their later educational attainment (Black *et al.* (2007)).¹²

Table 3.3 shows the coefficients on the instruments twins at last birth and the sex mix of the first two children. Both instruments have a significant effect on the number of children. The partial F-statistic on the instrument twins at last birth is equal to 55.50, and equal to 21.16 for the instrument same sex. An instrument should have a sufficiently strong effect on the endogenous explanatory variable and an often used rule of thumb is that the partial F-statistic should be larger than 10, which is the case for both instruments here.¹³

Table 3.3: OLS & IV results: effect of number of children on years of schooling of the first-born child

	OLS		IV		IV	
	(1)		Twins at last birth		Same sex	
	(1)		(2)		(3)	
Number of children	-0.02	(0.03)	0.30	(0.37)	0.59	(0.52)
Coefficient first stage			0.89***	(0.12)	0.11***	(0.02)
Partial F-statistic first stage			55.50		21.16	
Number of observations	5735		5735		5735	

All regressions include dummies for age and gender of the oldest child and controls for father's and mother's schooling. Standard errors are in parentheses. *significant at 10% level, **significant at 5% level, *** significant at 1% level

¹²Black *et al.* (2007) find using within twin techniques with data from Norway, that birth weight has a significant effect on educational attainment.

¹³Staiger and Stock (1997) investigate the finite sample bias of the IV estimator relative to the bias of the OLS estimator. They find that in a simple model the inverse of the F-statistic is an approximate estimate of the relative bias of the IV estimator. If the F-statistic is larger than 10, this gives a finite sample bias of the instrumental variable estimator of no more than 10 % of the OLS bias.

Table 3.3 also shows the OLS results for the first-born child, without taking into account the potential endogeneity of the number of children. The OLS coefficient in column 1 is negative but not significantly different from zero.¹⁴ The second stage estimation results in columns 2 and 3 show that the effect of the number of children on years of schooling of the oldest child is positive, but not significant. Although the standard errors are not small, this indicates that there is no significant negative impact of the number of children on years of schooling, because both second stage results give a positive coefficient on the number of children, irrespective of whether twins at last birth or the sex mix of the first two children is used as instrument.¹⁵ This result is similar to findings by recent studies using exogenous variation in the number of children to estimate the effect on years of schooling, in the sense that these studies also find no significant negative effect of family size (Black *et al.* (2005B) and Angrist *et al.* (2007)).

Birth order and educational attainment

Table 3.2 showed that including birth order dummies makes the coefficient on the number of children smaller and insignificantly different from zero, while all the birth order dummies are significantly negative. It could be that the negative coefficients on family size in columns 1 and 2 were just picking up the effect of birth order. To identify the effect of birth order, equation (3.4) will be estimated separately for families with two, three, four and five children, such that estimated effects cannot confound with the effect of family size, or with family characteristics which are correlated with family size.

$$E_{child} = \beta \cdot birthorder + X \cdot \phi + \eta \quad (3.4)$$

¹⁴The specification in column 1 in Table 3.3 is comparable to the specification in column 3 in Table 3.2. Column 3 in Table 3.2 controls for parental schooling and birth order and column 1 in Table 3.3 controls for parental schooling and is conditional on birth order since it only looks at first-born children.

¹⁵Due to the large standard errors it is however not possible to reject the hypothesis that the IV coefficients are equal to the OLS coefficient.

Table 3.4: The effect of birth order on child's years of schooling, by family size

	Two child family		Three child family		Four child family		Five child family	
	no	yes	no	yes	no	yes	no	yes
Birth order	-0.31*** (0.06)	-0.16 (0.15)	-0.40*** (0.04)	-0.26*** (0.07)	-0.36*** (0.04)	-0.27*** (0.06)	-0.29*** (0.05)	-0.20*** (0.08)
Second born	-0.31*** (0.06)		-0.44*** (0.06)		-0.45*** (0.08)		-0.49*** (0.13)	
Third born			-0.80*** (0.07)		-0.80*** (0.09)		-0.57*** (0.15)	
Fourth born					-1.10*** (0.11)		-1.10*** (0.16)	
Fifth born							-1.13*** (0.20)	
Family fixed effects	no	no	yes	no	no	no	no	no
Nr. families	2099	2091	2091	1262	538			
N	4020	5919	4699	2475				

All regressions include dummies for age and gender of the child. Standard errors (in parentheses) allow for correlation within families.
 * significant at 10% level, ** significant at 5% level, *** significant at 1% level.

Table 3.4 gives the results of estimating equation (3.4). For each family size the first column gives the estimates of including a set of dummy variables indicating whether the child is a second, third, fourth or fifth born, and X includes dummies for age and gender of the child.¹⁶ For all family sizes considered, the birth order dummies are all significantly negative, and the higher the birth order the more negative the effect is. The decline in years of schooling with birth order is approximately linear. Also the estimated birth order effects are very similar for the different family sizes. There is for example no large systematic difference between the effect of being a third born child in a three child family and being a third born in a four child family, whereas for a given family size there is a large difference between children with a different birth order.

Because birth order has an effect which is almost linear, the effect of birth order could as well be estimated by using a continuous variable, which has the value zero when an individual is a first-born, the value one when he or she is a second born etc. This gives a variable which ranges from zero to four. Results using this continuous variable instead of the birth order dummies are shown in the second columns in Table 3.4. These results also show that birth order has a relatively large negative effect on years of schooling. The coefficient is between 0.29 and 0.40 for all family sizes considered.

Since the WLS contains information about completed years of schooling for all the children within a family, it is possible to estimate equation (3.4) including family fixed effects. By including family fixed effects, birth order effects are identified on the basis of within family variation and it removes all observed and unobserved family characteristics that are constant over time. The results in Table 3.4 show that although the coefficients become somewhat smaller, including family fixed effects does not change the conclusion that child's years of schooling significantly declines with birth order.¹⁷

One possible reason for the negative effects is that the intellectual environment declines

¹⁶Later born children are younger, it is therefore important to include controls for birth cohort (age of the child), since educational opportunities might change over time.

¹⁷Only for two child families the coefficient is no longer significantly different from zero, which can be due to the large increase in the standard error.

with birth order, like Zajonc predicts in his confluence model. Another possibility is that earlier born children receive a larger share of parental time and resources compared to their later born siblings. The next section will investigate whether parental resources or competition between closely spaced siblings are behind the negative effect of birth order.

3.5 Potential mechanisms behind the birth order effects

The effect of child-spacing

Different theories predict that if families are more closely spaced, birth order effects will be larger. The confluence model from psychology states that the intellectual environment that later born children enter, will not differ much from the intellectual environment that earlier born children entered, when the age gap between earlier born and later born children is sufficiently large. When earlier born children are already almost adults when a child is born, this child does not suffer much from his higher birth order, in terms of a lower intellectual environment. This theory therefore predicts a decline in the birth order effects with the average space between births.

Also economic theories that stress the importance of resources predict that birth order effects will be smaller when the space between children is larger. If there are constraints on parental time this likely hurts later born children more than earlier born children, since later born children have to share the available time with their siblings for a larger part of their childhood. The constraints on the time that parents can spend with their children will however be less severe if the births of children are spread out over a longer time period. Birth order effects are therefore expected to be less negative when the average time between the births of the children is larger.

To test the effect of the average number of months between births, the following two

equations will be estimated.

$$E_{child} = \alpha_0 + \alpha_1 \cdot birthorder + \alpha_2 \cdot averagespace + X \cdot \beta + \varepsilon \quad (3.5)$$

$$E_{child} = \delta_0 + \delta_1 \cdot birthorder + \delta_2 \cdot averagespace + \delta_3 \cdot (averagespace \times birthorder) + X \cdot \gamma + \eta \quad (3.6)$$

The first equation includes the average number of months between births in a family of two, three, four or five children. The continuous birth order variable is included in the regression as well as years of schooling of the father and mother, and dummies for the age and gender of the child. If competition between closely spaced siblings has a negative effect on their educational attainment, the coefficient on the average space will likely be positive. Also the confluence model predicts a positive coefficient.

The second equation includes as an extra variable the interaction of birth order and the average number of months between births, so as to investigate how spacing affects the effect of birth order. If the negative effect of birth order is indeed smaller when the average time between births is larger, the coefficient on the interaction term (δ_3) will be positive. Table 3.5 shows the estimation results of the two equations.

The average number of months between two births does not have a significant effect on the educational attainment of a child, as can be seen in Table 3.5. Also the interaction terms of birth order and the average space are insignificant and this holds for all family sizes. These results indicate that competition between closely spaced siblings does not seem to be an important cause of the negative effect of birth order.

The time between subsequent births can however be chosen by the parents. If certain characteristics of parents have an effect on the average space between births, as well as on the education of the children, the average number of months between births is endogenous. The last columns in Table 3.5 estimate equation (3.6) including family fixed effects, to control for parental characteristics that might be correlated with child-spacing and that are

Table 3.5: Effect of the average space, birth order and the interaction of the two on child's years of schooling

	Two child family		Three child family		Four child family		Five child family					
(Average space × birth order)	0.000 (0.003)	-0.013 (0.016)	-0.001 (0.002)	0.000 (0.004)	0.003 (0.002)	-0.003 (0.003)	-0.001 (0.004)	0.002 (0.005)				
Average space	-0.003 (0.002)	-0.003 (0.002)	0.001 (0.002)	-0.000 (0.003)	-0.002 (0.004)	-0.005 (0.005)	-0.011 (0.007)	-0.009 (0.010)				
Birth order	-0.15** (0.06)	-0.15 (0.12)	-0.11 (0.17)	-0.24*** (0.04)	-0.29*** (0.07)	-0.24** (0.09)	-0.21*** (0.04)	-0.29*** (0.07)	-0.27*** (0.08)	-0.17*** (0.05)	-0.13 (0.11)	-0.17 (0.13)
Family fixed effects	no	no	yes	no	no	yes	no	no	no	no	yes	
Nr. families	2013	1936	1936	5502	4209	2121	457	457	2121	2121	2121	
N	3861	5502	4209	2121	457	2121	457	2121	457	2121	457	

All regressions include dummies for age and gender of the child and controls for father's and mother's years of schooling. Standard errors (in parentheses) allow for correlation within families. * significant at 10% level, ** significant at 5% level, *** significant at 1% level.

fixed over time. The results show that including family fixed effects does not change the results, the coefficients on the interaction between birth order and the average space are very small and none of them is significantly different from zero.

The fixed effect specification only controls for parental characteristics that do not change over time, and therefore does not fully solve the potential endogeneity of the spacing between births. To identify the causal effect of the average time between births, the presence of twins among the children in a family is used as an instrument. If a twin is born the space between two births is zero, which is an exogenous decrease in the average time between births. The following two equations will be estimated

$$E_{child} = \alpha_0 + \alpha_1 \cdot averagespace + X \cdot \beta + v \quad (3.7)$$

$$E_{firstborn} - E_{lastborn} = \delta_0 + \delta_1 \cdot averagespace + X \cdot \phi + v \quad (3.8)$$

Whereby the average space is instrumented by a variable which takes on the value one when either the first, second, third or fourth birth is a twin.

A twin birth also has a significant effect on the number of children as was shown in the previous section. The equations above will however be estimated separately by family size, so this will not be an issue here. Another potential problem is that if the education of a child is affected by being part of a twin the instrument could have an effect on years of schooling separate from the effect via the average space.¹⁸ For children from families with at least three children it is possible though to estimate equation (3.7) excluding the children who are part of a twin.

Table 3.6 shows the effect of the average number of months between births on child's years of schooling, as well as the coefficients from the first stage and the partial F-statistics. The presence of twins has a significant negative effect on the average number of months between births. There is no effect though of the average space on child's years of school-

¹⁸As was already mentioned in previous sections, children part of a twin have on average a lower birth weight compared to singletons and this might affect there later educational attainment.

ing, except for four child families, but this is only significant at the margin. These findings confirm the results in Table 3.5.

Table 3.6: Effect of the average space on years of schooling, instrumented by the presence of twins

	Three child family		Four child family		Five child family	
Average space	-0.03	(0.02)	-0.05*	(0.02)	-0.03	(0.04)
Coefficient first stage	-12.32***	(2.57)	-9.34***	(1.06)	-6.54***	(0.95)
Partial F-statistic first stage	23.04		77.62		47.20	
Number of observations	5434		4103		2059	

Excluding children who are part of a twin. Regression includes dummies for age and gender of the child, father's and mother's years of schooling and birth order. Standard errors (in parentheses) control for correlation within families. *significant at 10% level, **significant at 5% level, ***significant at 1% level

Table 3.7: Effect of the average space on schooling difference first and last born child instrumented by presence of twins

	Three child family		Four child family		Five child family	
Average space	0.01	(0.03)	0.09*	(0.05)	-0.12	(0.12)
Coefficient first stage	-11.85***	(2.70)	-8.04***	(1.61)	-5.28***	(1.83)
Partial F-statistic first stage	19.27		24.80		8.29	
Number of observations	1718		1020		402	

Regression includes control variables for gender of the first and last born child and father's and mother's years of schooling. Standard errors are in parentheses *significant at 10% level, **significant at 5% level, ***significant at 1% level

To investigate whether child-spacing affects the difference in schooling between earlier and later born children, Table 3.7 shows the results of estimating equation (3.8). Sample sizes in Table 3.7 are smaller because it looks at the difference between the first and the last born child and therefore uses only one observation per family. The second stage results

show that if the time between births is larger this does not change birth order effects, since the average space has no significant effect on the schooling difference between the first and the last born child.¹⁹

The allocation of parental resources

Although economic theories, that stress the importance of parental resources, predict birth order effects to decline with the average space between births, the finding that spacing does not affect the birth order patterns does not rule out that the allocation of time and other parental resources is a potential mechanism behind the birth order effects. Price (2008) finds for example that parents spend equal amounts of time with their children at a particular point in time, but they spend less time with their children when the first-born child becomes older. This hurts later born children, because at a particular age later born children receive less quality time compared to the amount of quality time received by the first-born child at that same age. This could be an explanation for the negative effect of birth order on children's schooling outcomes.

The Wisconsin Longitudinal Study does not contain information about the amount of time parents spend with their children. It does however contain information about financial transfers to children. In 1992 respondents (both the graduates and the selected siblings) were asked whether they or their spouse had given or loaned any of their children \$1000 or more between 1975 and 1992. The questionnaire contained seven different reasons for giving money; educational expenses, down payment on a home, enter or continue business or farm, increase wealth or decrease debt, medical expenses, housing or other living expenses, and other. For each reason the respondents were asked whether they or their spouse had given or loaned money for that reason to any of their children. If the respondent answered yes the next question was which child received the most money, second most, third most etc. For any child that received money the total amount received

¹⁹Except for four child families, but again only significant at the margin.

was determined. Parents could report that they gave equal amounts to their children. In that case the amount given to the child who received the most is equal to the amount given to the child who received the second most etc.

The WLS data set does not contain the direct answers to the questions described above. Instead it contains variables with the birth order of the child who received the most, second most, third most etc, variables with the amount received by the child who received the most, second most, etc and variables with the main reason for receiving money. These variables were created by matching the name of the child who received money to the same name in the children's roster. This children's roster contains characteristics of all of the children of the respondent, such as their birth order, gender and years of schooling.²⁰ Based on these variables it is possible to determine for each child whether it received any money from his parents between 1975 and 1992, and if so the amount he or she received.

Less than 2% of the respondents refused to answer questions about giving to children. About 23% answered that they did not give any money to their children and another 7% reported giving money to at least one child but responded with don't know or refused to all questions about the amounts given. Only adult children could have possibly received money from their parents because of the reasons for giving described above. The analysis below will therefore use a sample of families in which all children were older than 25 in 1992, such that all children within a family had the possibility of receiving money from their parents.

The analysis below will look at two variables. The first variable takes on the value one when a child received \$1000 or more from his parents and takes on the value zero otherwise. The second variable contains the amount. This variable is defined for families in which at least one child received money and is equal to \$1000 or more if the child received money from his parents.

²⁰In 90% of the cases a computer program was able to match the name and another 7% were matched by hand-coding, the remaining 3% could not be matched because the name was not in the roster or because of a response such as "son" when there was more than one son etc.

Before investigating whether the financial transfers are affected by the number of children or the birth order of the children, Table 3.8 estimates the relation between receiving money from parents and child's years of schooling.

Table 3.8: Regressing child's years of schooling on probability of receiving money/amount received

	Two child family		Three child family		Four child family		Five child family	
Received any money	1.40*** (0.13)		1.44*** (0.11)		1.46*** (0.14)		1.29*** (0.22)	
Ln(amount)		0.18*** (0.02)		0.16*** (0.02)		0.14*** (0.02)		0.09*** (0.03)
<i>N</i>	1583	1203	2160	1582	1349	876	603	398

All regressions include dummies for age and gender of the child. Standard errors (in parentheses) allow for correlation within families. *significant at 10% level, **significant at 5% level, ***significant at 1% level

The results show that children who receive \$1000 or more from their parents have a significantly higher educational attainment. To see whether the amount they receive is important, Table 3.8 also includes a specification with the natural logarithm of the amount received as explanatory variable.²¹ For all family sizes considered the logarithm of the amount received is significantly related to child's years of schooling. Although the results do not necessarily reflect a causal impact of receiving money on educational outcomes, it does indicate that if receiving money differs by birth order, this could be a potential mechanism behind the birth order effects, especially since for about 60% of the children the main reason for receiving money was to cover their educational expenses.

The quantity-quality model predicts that the investment in child quality will be lower when there are more children. To investigate whether the number of children indeed negatively affects the probability that the first-born child receives money, Table 3.9 shows the effect of estimating equation (3.9), whereby the dependent variable takes on the value

²¹This variable is created by taking the logarithm of the amount plus 1, so as to account for the fact that some children received an amount of zero.

one when the first-born child received \$1000 or more from his parents, and X includes dummies for age and gender of the child.

$$Anymoney = \alpha + \beta \cdot children + X \cdot \phi + \varepsilon \quad (3.9)$$

Column 1 of Table 3.9 shows that the probability that the first-born child receives any money is significantly negatively related to the number of children. To account for the endogeneity of family size, column 2 shows the results when the number of children is instrumented by both twins at last birth and the sex composition of the first two children. The IV estimate is lower than the OLS estimate and is no longer significantly different from zero, indicating that neither the investment in children nor their completed schooling outcomes are significantly affected by an increase in the number of children.²²

Table 3.9: Effect of the number of children on the probability that first-born child receives money

	OLS (1)	IV (2)
Number of children	-0.07*** (0.01)	-0.03 (0.11)
Coefficients first stage:		
twins at last birth		0.97*** (0.19)
same sex		0.10*** (0.04)
Partial F-statistic first stage		14.65
Number of observations	1973	1973

All regressions include dummies for age and gender of the first-born child. Standard errors are in parentheses. * significant at 10% level, ** significant at 5% level, *** significant at 1% level

To investigate whether receiving money varies with the birth order of the child, the

²²Due to the large standard errors it is however not possible to reject the hypothesis that the IV coefficients are equal to the OLS coefficient.

following two equations will be estimated separately for different family sizes.

$$\text{Anymoney} = \alpha + \beta \cdot \text{birthorder} + X \cdot \phi + \varepsilon \quad (3.10)$$

$$\ln \text{amount} = \gamma + \delta \cdot \text{birthorder} + X \cdot \pi + \eta \quad (3.11)$$

Table 3.10 shows the results of estimating equation (3.10). The results show that the probability of receiving money significantly declines with the birth order of the child. Only for five child families the coefficient is not significantly different from zero.

Table 3.10: Effect of birth order on probability of receiving any money from parents

	Two child family	Three child family	Four child family	Five child family
Birth order	-0.08*** (0.02)	-0.04** (0.02)	-0.04** (0.02)	0.01 (0.02)
<i>N</i>	1583	2160	1349	603

Regression includes dummies for age and gender of the child. Standard errors (in parentheses) control for correlation within families. *significant at 10% level, **significant at 5% level, ***significant at 1% level

Table 3.11 estimates the effect of birth order on the logarithm of the amount received for families in which at least one child receives money. To account for the fact the amount is a censored variable, since the amount is only specified for children that received \$1000 or more, the censored least absolute deviations (CLAD) estimator introduced by Powell (1984) is used to estimate equation (3.11). The CLAD estimator, or conditional median regression, is a semi-parametric method to account for censored observations and can be used if less than 50% of the observations are censored, because then the censored sample median provides a consistent estimate of the population median. Since for five child families less than half of the observations received an amount of \$1000 or more, only two, three and four child families are investigated.

Table 3.11: Effect of birth order on the logarithm of the amount received from parents (CLAD estimator)

	Two child family	Three child family	Four child family
Birth order	-0.36*** (0.14)	-0.34** (0.15)	-0.51* (0.30)
Number of observations	1203	1582	876

Regression includes controls for age and gender of the child. Standard errors (in parentheses) control for correlation within families. *significant at 10% level, **significant at 5% level, ***significant at 1% level

The results in Table 3.11 show that not only the probability to receive any money from parents is higher for earlier born children but also the amount received varies with the birth order of the child. Later born children receive significantly less money than earlier born children. These results indicate that restrictions on parental resources are a likely mechanism behind the negative birth order effects that are found in this and in other studies.

3.6 Conclusion

This chapter has investigated the causal effect of the number of siblings and birth order on years of schooling. By using twins at last birth and the sex mix of the first two children as instruments, the effect of the number of children on the educational attainment of the oldest child is identified. Although the OLS estimate of the effect of the number of children on child's years of schooling is negative, the IV results show a positive and insignificant effect of the number of children on years of schooling of the oldest child. This indicates that the negative correlation between sibship size and educational attainment found in this chapter, and found in other papers (Belmont and Marolla (1973), Blake (1981), Hanushek (1992)), might actually be caused by unobserved family characteristics, which affect the number of children as well as the educational achievements of those children.

The effect of birth order is identified by investigating the effect of birth order for two, three, four and five child families separately. Birth order turns out to have a significant negative effect on years of education, for all family sizes examined. This is in contrast to the findings by Blake (1981), Hauser and Sewel (1985) and Hanushek (1992) but the results are very similar to those found by Behrman and Taubman (1986), Black, Devereux and Salvanes (2005B) and Kantarevics and Mechoulan (2006).

Although a number of studies have investigated the impact of birth order and family size on educational outcomes, very few papers have studied potential mechanisms behind these effects. This chapter has looked at two potential mechanisms, competition between closely spaced siblings and the allocation of parental resources to children. Child-spacing, measured by the average number of months between births, does not affect the birth order patterns that are observed in children's schooling outcomes. This results is robust to including family fixed effect and the use of the presence of twins as instrument.

The results on financial transfers to children show a negative birth order pattern just as the results found on children's schooling outcomes. The differential allocation of resources to earlier and later born children could therefore be a reason for the negative effect of birth order on child's years of schooling. But why would parents give less money to later born children? One possible explanation is consistent with the resource dilution idea. If parents are not enough forward looking or if they have difficulties in estimating the amount of money necessary to invest in their children's schooling, it could be that later born children will find that most parental resources have been used by earlier born siblings at the moment they need money to go to college.

Another explanation could be that earlier born children already receive a larger share of parental time and resources (such as a bigger room) when they are young. This could make earlier born children do better in school and if parents invest in the children that perform best in school, this might be a reason why earlier born children have a higher probability of receiving money from their parents.

Both explanations suggest that parental resources are a likely mechanism behind the birth order effects. However, more research is necessary to get a better understanding of why parents tend to allocate more resources to earlier born children. Also information about resource allocation and the effect on schooling outcomes when children are still quite young would provide more insight into the exact mechanisms behind the estimated birth order effects.

Chapter 4

Summary and conclusions

The chapters in this thesis investigate the effect of father's and mother's schooling, the number of children and birth order on children's completed years of schooling. All the results in the chapters are based on the Wisconsin Longitudinal Study, which is a long term study of 10,317 high school graduates from Wisconsin high schools in 1957. These graduates have been followed over time and information has been collected on many topics including schooling outcomes and family characteristics, which makes it a very suitable data set to investigate the impact of family background on children's schooling outcomes.

Chapter 1 investigates the causal impact of father's and mother's schooling on child's schooling. Previous studies have used different identification strategies to estimate the causal impact of parents' schooling. These identification approaches generally put strong requirements on the data and give point estimates which are only informative if the assumptions on which they are based are correct. This chapter uses a different approach to investigate the effect of parents' schooling on child's schooling, a nonparametric bounds analysis. By making relatively weak and testable assumptions this chapter obtains bounds on the effect of increasing parents' schooling on years of schooling of the child and on the probability that the child obtains a bachelor's degree. These bounds are obtained by subsequently adding a monotone treatment response assumption (MTR), a monotone treatment selection assumption (MTS) and a monotone instrumental variable assumption

(MIV), whereby the schooling of the grandparent and the schooling of the spouse are used as monotone instrumental variables.

Although the bounds on the effect of parents' schooling include a zero effect, the analysis obtains informative upper bounds especially for the effect of increasing parents' schooling from a high school degree to a bachelor's degree. Both for mothers as for fathers the MTR-MTS-MIV upper bounds are significantly lower than the OLS estimates. These results show that the effect of parents' schooling on child's schooling is lower than what one would conclude on the basis of simple correlations.

Chapter 2 studies the problem that recent studies that make a distinction between causation and selection when estimating intergenerational schooling mobility, often rely on samples in which information on the child's completed schooling is not always available. This chapter estimates the impact of mother's and father's schooling on child's schooling using censored and uncensored samples of own birth children and adoptees, and investigates the consequences of three different methods that deal with censored observations: maximum likelihood approach, replacement of observed with expected years of schooling and elimination of all school-aged children.

Using uncensored samples, this chapter shows that the estimate of the relation between mother's schooling and child's schooling is significantly smaller when estimated on a sample of adoptees instead of on a sample of own birth children, while the effect of father's schooling is very similar in both samples. Using a sample in which part of the children is still in school gives a small downward bias in the estimates. When the problem of censored observations is tackled by either three correction methods the estimates are no longer downwardly biased. Instead the maximum likelihood approach and the method whereby all school-aged children are eliminated both give a small but significant positive bias in the estimates. The method that treats parental expectations as if they were realizations performs best. This result depends, however, on the degree of censoring. For samples that are largely incomplete the method does give a small (negative) bias. The results suggest

that it doesn't matter (much) whether researchers are patient or impatient: whether we fully rely on parental expectations, or whether we use realizations measured 30 years later, the mobility estimates are not substantially different.

Chapter 3 investigates the impact of family size and birth order on years of schooling. By using twins at last birth and the sex mix of the first two children as instruments, the effect of the number of children on the educational attainment of the oldest child is identified. Although the OLS estimate of the effect of the number of children on child's years of schooling is negative, the IV results show a positive and insignificant effect of the number of children on years of schooling of the oldest child. The effect of birth order is identified by investigating the effect of birth order for two, three, four and five child families separately. Birth order turns out to have a significant negative effect on years of education, for all family sizes examined.

To investigate what is behind the negative birth order effects, this chapter looks at two potential mechanisms, competition between closely spaced siblings and the allocation of parental resources to children. Child-spacing, measured by the average number of months between births, does not affect the birth order patterns that are observed in children's schooling outcomes. This results is robust to including family fixed effect and the use of the presence of twins as instrument. The results on financial transfers to children show a negative birth order pattern just as the results found on children's schooling outcomes. These results suggest that the differential allocation of resources to earlier and later born children is a potential mechanism behind the negative effect of birth order on child's years of schooling.

The main conclusion that emerges from the results in these chapters is that family background is important for children's schooling outcomes, but that simply looking at the association between a specific family characteristic and children's outcomes is not sufficient and can lead to wrong conclusions. The family is a complex structure of various family characteristics and it is important to identify the causal impacts of the different

family background components. Also the method to estimate these effects should be chosen with care since it can generate biased estimates, when the assumptions on which the method is based turn out to be incorrect.

Bibliography

Angrist, J.D. and W.N. Evans (1998), 'Children and their parent's labor supply: Evidence from exogenous variation in family size', *American Economic Review* 88(3), 450–477.

Angrist, J.D., V. Lavy and A. Schlosser (2007), 'Multiple experiments for the causal link between the quantity and quality of children', *mimeo, Hebrew University, Department of Economics*.

Antonovics, K.L. and A.S. Goldberger (2005), 'Does increasing women's schooling raise the schooling of the next generation? Comment', *American Economic Review* 95(5), 1738–1744.

Arabmazar, A. and P. Schmidt (1982), 'An investigation of the robustness of the tobit estimator to non-normality', *Econometrica* 50(4), 1055–1064.

Becker, G.S. (1981), *A treatise on the family*, Cambridge: Harvard University Press.

Becker, G.S. and H.G. Lewis (1973), 'On the interaction between the quantity and quality of children', *Journal of Political Economy* 81(2), S279–S288.

Becker, G.S. and N. Tomes (1976), 'Child endowments and the quantity and quality of children', *Journal of Political Economy* 84(4), S143–S162.

Becker, G.S. and N. Tomes (1979), 'An equilibrium theory of the distribution of income and intergenerational mobility', *Journal of Political Economy* 87(6), 1153–1189.

Behrman, J. (1997), 'Intrahousehold distribution and the family', in Rosenzweig M. R. and O. Stark, eds, *Handbook of Population Economics*, Vol. I, Amsterdam: Elsevier Science.

Behrman, J.R. and M.R. Rosenzweig (2002), 'Does increasing women's schooling raise the schooling of the next generation?', *American Economic Review* 92(1), 323–334.

Behrman, J.R. and M.R. Rosenzweig (2005), 'Does increasing women's schooling raise the schooling of the next generation? Reply', *American Economic Review* 95(5), 1745–1751.

Behrman, J.R. and P. Taubman (1986), 'Birth order, schooling and earnings', *Journal of Labor Economics* 4(3), 121–145.

Belmont, L. and F. A. Marolla (1973), 'Birth order, family size, and intelligence', *Science* 182(4117), 1096–1101.

Ben-Porath, Y. and F. Welch (1976), 'Do sex preferences really matter?', *Quarterly Journal of Economics* 90(2), 285–307.

Björklund, A., M. Lindahl and E.J.S. Plug (2006), 'The origins of intergenerational associations: Lessons from swedish adoption data', *Quarterly Journal of Economics* 121(3), 999-1028.

Black, S.E., P.J. Devereux and K.G. Salvanes (2005A), 'Why the apple doesn't fall far: Understanding intergenerational transmission of human capital', *American Economic Review* 95(1), 437–449.

Black, S.E., P.J. Devereux and K.G. Salvanes (2005B), 'The more the merrier? The effect of family size and birth order on children's education', *Quarterly Journal of Economics* 120(2), 669–700.

Black, S.E., P.J. Devereux and K.G. Salvanes (2007), 'From the cradle to the labor market? The effect of birth weight on adult outcomes', *Quarterly Journal of Economics* 122(1), 409–439.

Blake, J. (1981), 'Family size and the quality of children', *Demography* 18(4), 421–442.

Blundell, R., A. Gosling, H. Ichimura and C. Meghir (2007), 'Changes in the distribution of male and female wages accounting for employment composition using bounds', *Econometrica* 75(2), 323–363.

Böhlmark, A. and M.J. Lindquist (2006), 'Life-cycle variations in the association between current and lifetime income: Replication and extension for Sweden', *Journal of Labor Economics* 24(4), 879–896.

Bound, J., D.A. Jaeger and R.M. Baker (1995), 'Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak', *Journal of the American Statistical Association* 90(430), 443–450.

- Bound, J. and G. Solon (1999), 'Double trouble: On the value of twins-based estimation of the return to schooling', *Economics of Education Review* 18(2), 169–182.
- Butcher, K.F and A. Case (1994), 'The effect of sibling sex composition on women's education and earnings', *Quarterly Journal of Economics* 109, 531–563.
- Cáceres-Delpiano, J. (2006), 'The impacts of family size on investment in child quality', *Journal of Human Resources* 41(4), 738–754.
- Carneiro, P., C. Meghir and M. Parey (2007), 'Maternal education, home environments and the development of children and adolescents', *IZA Working Paper*, No. 3072.
- Chevalier, A. (2004), 'Parental education and child's education: A natural experiment', *IZA Working Paper*, No. 1153.
- Conley, D. and R. Glauber (2006), 'Parental educational investment and children's academic risk: Estimates of the impact of sibship size and birth order from exogenous variation in fertility', *Journal of Human Resources* 41(4), 722–737.
- Cunha, F., J.J. Heckman, L. Lochner and D.V. Masterov (2006), 'Interpreting the evidence on life cycle skill formation', in Hanushek, E. and F. Welch, eds, *Handbook of the Economics of Education*, Vol. I. Amsterdam: Elsevier Science.
- Cunha, F. and J.J. Heckman (2007), 'The technology of skill formation', *American Economic Review* 97(2), 31–47.
- Currie, J. and E. Moretti (2003), 'Mother's education and the intergenerational transmis-

sion of human capital: Evidence from college openings', *Quarterly Journal of Economics* 118(4), 1495–1532.

Gerfin, M., and M. Schellhorn (2006), 'Nonparametric bounds on the effect of deductibles in health care insurance on health care demand – swiss evidence', *Health Economics* 15(9), 1011-1020.

González, L. (2005), 'Nonparametric bounds on the returns to language skills', *Journal of Applied Econometrics* 20, pp. 771-795.

Haider, S.J. and G. Solon (2006), 'Life-cycle variation in the association between current and lifetime earnings', *American Economic Review* 96(4), 1308–1320.

Hanushek, E.A. (1992). 'The trade-off between child quantity and quality', *Journal of Political Economy* 100(1), 84–117.

Hauser, R.M. and H.D. Kuo (1998). 'Does the gender composition of sibships affect women's educational attainment?', *Journal of Human Resources* 33(3), 644–657.

Hauser, R.M., and W.H. Sewel (1985), 'Birth order and educational attainment in full sibships', *American Educational Research Journal* 22(1), 1–23.

Haveman, R., and B. Wolfe (1995), 'The determinants of children attainments: A review of methods and findings', *Journal of Economic Literature* 33(4), 829–78.

Hertz, T., T. Jayasundera, P. Piraino, S. Selcuk, N. Smith and A. Verashchagina (2007), 'The inheritance of educational inequality: International comparisons and fifty-year trends',

The B.E. Journal of Economic Analysis & Policy 7(2).

Available at: <http://www.bepress.com/bejeap/vol7/iss2/art10>

Kaestner, R. (1997), 'Are brothers really better: Sibling sex composition and educational achievement revisited', *Journal of Human Resources* 32(2), 250–284.

Kantarevic, J. and S. Mechoulan (2006), 'Birth order, educational attainment, and earnings: An investigation using the PSID', *Journal of Human Resources* 41(4), 755–777.

Lechner, M. (1999), 'Nonparametric bounds on employment and income effects of continuous vocational training in East Germany', *Econometrics Journal* 2, 1–28.

Manski, C.F. (1989), 'Anatomy of the selection problem', *Journal of Human Resources* 24(3), 343–360.

Manski, C.F. (1997), 'Monotone treatment response', *Econometrica* 65(6), 1311–1334.

Manski, C.F., and J.V. Pepper (2000), 'Monotone instrumental variables: With an application to the returns to schooling', *Econometrica* 68, 997–1010.

Maurin, E. and S. McNally (2008), 'Vive la révolution ! Long term returns of 1968 to the angry students', *Journal of Labor Economics* 26(1), 1–35

Oreopoulos, P., M.E.J. Page and A.H. Stevens (2006), 'The intergenerational effects of compulsory schooling', *Journal of Labor Economics* 24(4), 729–760.

Organisation for Economic Co-operation and Development—Meeting social policy min-

isters (2005), www.oecd.org/socialmin2005, URL accessed on July 10, 2008.

Pepper, J.V. (2000), 'The intergenerational transmission of welfare receipt: A nonparametric bounds analysis', *Review of Economics and Statistics* 82(3), 472–478.

Plug, E.J.S. (2004), 'Estimating the effect of mother's schooling on children's schooling using a sample of adoptees', *American Economic Review* 94, 358–368.

Price, J. (2008), 'Parent-child quality time: Does birth order matter?', *Journal of Human Resources* 43(1), 240–265.

Rosenzweig, M.R., and K.I. Wolpin (2000), 'Natural "natural experiments" in economics', *Journal of Economic Literature* 38(4), 827–874.

Sacerdote, B.I. (2002), 'The nature and nurture of economic outcomes', *American Economic Review* 92(2), 344–348.

Sacerdote, B.I. (2007), 'How large are the effects from changes in family environment? A study of korean american adoptees', *Quarterly Journal of Economics* 122(1), 119–157.

Sewell, W.H., R.M. Hauser, K. Springer and T.S. Hauser (2004), *As we age: The Wisconsin Longitudinal Study, 1957-2001*. London: Elsevier Ltd.

Solon, G. (1999), 'Intergenerational mobility in the labor market', in Ashenfelter, O. and S. Card, eds, *Handbook of Labor Economics*, Vol. III. Amsterdam: Elsevier Science.

Staiger, D., and J.H. Stock (1997), 'Instrumental variables regression with weak instru-

ments', *Econometrica* 65(3), 557–586.

U.S. Network for Education Information: U.S. Department of Education,
<http://www.ed.gov/about/offices/list/ous/international/usnei/us/edlite-struct-geninfo.html>,
URL accessed on March 24, 2008.

Wisconsin Longitudinal Study (WLS) [graduates and siblings]: 2004/2005 Version 12.20.
[machine-readable data file] / Hauser, Robert M; Hauser, Taissa S. [principal investigator(s)].
Madison, WI: University of Wisconsin-Madison, WLS. [distributor];
<http://www.ssc.wisc.edu/wlsresearch/documentation/>.

Wisconsin Longitudinal Study (WLS) (2006). *Wisconsin Longitudinal Study handbook*,
E. Wolmering, eds, with Contributions by Wisconsin Longitudinal Study Staff.

Zajonc, R.B. (1976), 'Family configuration and intelligence', *Science* 192, 227–236.

Nederlandse samenvatting

Wat bepaalt het succes van kinderen? Zowel wetenschappers als beleidsmakers onderstrepen het belang van de familie voor de latere socio-economische uitkomsten van kinderen. Er zijn meerdere familiekenmerken die van invloed kunnen zijn op de uitkomsten van kinderen en deze zijn vaak sterk aan elkaar gerelateerd. Het is daarom belangrijk om de effecten van de verschillende kenmerken uit elkaar te halen. Als een associatie wordt geïnterpreteerd als een causale relatie dan kan dit namelijk leiden tot verkeerde conclusies. Een familiekenmerk kan dan een effect lijken te hebben op de socio-economische uitkomsten van een kind, terwijl het eigenlijk het effect oppikt van iets anders.

Dit proefschrift richt zich op één socio-economische uitkomst; de hoogst behaalde opleiding van het kind. In de hoofdstukken van dit proefschrift wordt onderzocht wat het effect is van de opleiding van de vader en moeder, gezinsgrootte en geboortevolgorde op het aantal afgeronde jaren onderwijs van kinderen. Alle resultaten in de hoofdstukken zijn gebaseerd op de Wisconsin Longitudinal Study (WLS). Dit is een lange termijn studie van 10,317 afgestudeerden van middelbare scholen in Wisconsin (United States) in 1957. Deze afgestudeerden zijn gevolgd in de tijd en informatie is verzameld over tal van onderwerpen, onder andere over onderwijsuitkomsten en over verscheidene familiekenmerken.

Hoofdstuk 1 onderzoekt het causale effect van de opleiding van de vader en moeder op de opleiding van het kind. Eerdere studies hebben gebruik gemaakt van verschillende strategieën om het causale effect van de opleiding van de ouders te identificeren. Deze identificatiemethoden eisen vaak veel van de data en geven puntschattingen die alleen informatief zijn als de veronderstellingen waarop zij zijn gebaseerd juist zijn. Dit hoofdstuk

gebruikt een relatief nieuwe aanpak, een niet-parametrische begrenzingmethode, voor het onderzoek naar de invloed van de opleiding van de ouders op de opleiding van kinderen. Deze methode schat intervallen waarbinnen een bepaald effect valt. De analyse begint zonder enige veronderstellingen en vervolgens worden er relatieve milde en testbare veronderstellingen toegevoegd die de intervallen kleiner maken. Het hoofdstuk schat boven- en ondergrenzen op het effect van een toename in de opleiding van de ouders op het aantal opleidingsjaren van het kind en op de kans dat het kind een Bachelors graad behaalt.

De geschatte intervallen in dit hoofdstuk kunnen een nul effect niet uitsluiten. De bovengrenzen zijn echter wel informatief, vooral voor het effect van een toename in de opleiding van de ouders van een middelbare school diploma tot een Bachelors graad. De niet-parametrische bovengrenzen zijn zowel voor vaders als voor moeders significant lager dan de resultaten die verkregen zijn met de kleinste-kwadraten methode (ook wel OLS genoemd). Deze bevindingen laten zien dat het causale effect van de opleiding van de ouders op de opleiding van het kind lager is dan wat we zouden concluderen op basis van eenvoudige correlaties.

Hoofdstuk 2 schat het effect van de opleiding van de moeder en de vader op de opleiding van het kind en maakt daarbij gebruik van steekproeven van biologische en geadopteerde kinderen. De ouders en kinderen zijn gevolgd over de tijd en er is informatie verzameld over de opleidingsuitkomsten van de kinderen in 1992 en in 2004. In 1992 zit een deel van de kinderen nog op school en in 2004 hebben alle kinderen hun opleiding afgerond. Het hoofdstuk onderzoekt het probleem van het gebruik van een steekproef waarin een deel van de kinderen nog op school zit en evalueert drie verschillende oplossingen voor dit probleem: de methode van de meest aannemelijke schatter (ook wel maximum likelihood methode genoemd); de vervanging van waargenomen door verwachte jaren opleiding; en het verwijderen van alle schoolgaande kinderen uit de steekproef.

Wanneer gebruik wordt gemaakt van de meest recente steekproef dan blijkt dat de relatie tussen de opleiding van moeder en kind aanzienlijk kleiner is wanneer deze relatie

geschat wordt op een steekproef van geadopteerde kinderen in plaats van op een steekproef van biologische kinderen. De relatie tussen de opleiding van vader en kind is daarentegen ongeveer gelijk in beide steekproeven.

Wanneer een steekproef wordt gebruikt waarbij een deel van de kinderen nog op school zit, dan geeft dit een onderschatting van de relatie tussen de opleiding van ouders en hun kind. Als één van de drie bovengenoemde methoden wordt toegepast, dan verdwijnt deze onderschatting. In plaats daarvan geven de maximum likelihood methode en de methode waarbij alle schoolgaande kinderen uit de steekproef worden verwijderd, beide een kleine maar significante overschatting. De methode waarbij de waargenomen, niet afgeronde, opleiding wordt vervangen door de verwachtingen van ouders presteert het beste. Dit resultaat hangt echter af van het percentage kinderen in de steekproef dat nog op school zit. Als bijna alle kinderen in de steekproef hun opleiding nog niet hebben afgerond, dan geeft deze methode wel een kleine onderschatting.

De resultaten in dit hoofdstuk duiden erop dat het niet zoveel uitmaakt of onderzoekers heel ongeduldig zijn en een steekproef gebruiken waarin bijna alle kinderen nog op school zitten of dat ze wachten totdat alle kinderen hun opleiding hebben afgerond. Als de relatie tussen de opleiding van ouders en hun kind wordt geschat door gebruik te maken van verwachte scholingsuitkomsten dan geeft dit resultaten die niet veel verschillen van de resultaten wanneer we de werkelijke afgeronde opleidingsuitkomsten van het kind als afhankelijke variabele gebruiken.

Hoofdstuk 3 onderzoekt de effecten van gezinsgrootte en geboortevolgorde op het aantal afgeronde opleidingsjaren van een kind. Het effect van het aantal kinderen in een gezin op de opleidingsuitkomst van het oudste kind wordt geïdentificeerd door gebruik te maken de instrumentele variabele methode. Er worden twee instrumenten gebruikt voor gezinsgrootte; tweelingen bij laatste geboorte en de geslachtscompositie van de eerste twee kinderen. De resultaten die verkregen zijn met de kleinste-kwadraten methode duiden op een significante negatieve relatie tussen gezinsgrootte en de opleidingsuitkom-

sten van de kinderen. De resultaten die verkregen zijn door gebruik te maken van de instrumentele variabele methode zijn daarentegen positief en niet significant verschillend van nul. Deze resultaten geven aan dat er geen significant negatief causaal effect is van het aantal kinderen op het aantal jaren opleiding van het oudste kind.

Het effect van geboortevolgorde wordt geïdentificeerd door het effect apart te schatten voor gezinnen met twee, drie, vier of vijf kinderen. Op deze manier kan geboortevolgorde niet het effect oppikken van gezinsgrootte of familiekenmerken die gerelateerd zijn aan gezinsgrootte. De resultaten laten zien dat geboortevolgorde een significant negatief effect heeft op de scholingsuitkomsten van kinderen.

Het hoofdstuk onderzoekt twee mogelijk mechanismen achter deze negatieve geboortevolgorde effecten; competitie tussen kinderen in gezinnen waar de tijdsafstand tussen de geboortes klein is en de verdeling van financiële middelen tussen de kinderen. De gemiddelde afstand tussen de geboortes heeft geen effect op de geboortevolgorde patronen. Dit resultaat verandert niet als we gebruik maken van gezinspecifieke effecten of als de aanwezigheid van een tweeling in het gezin wordt gebruikt als instrument voor de gemiddelde afstand tussen geboortes. Informatie over de hoeveelheid geld die ouders hebben gegeven aan hun kinderen laat zien dat eerder geboren kinderen een grotere kans hebben om geld te ontvangen van hun ouders dan hun jongere broers en zussen. Ook het bedrag dat ze krijgen is hoger. Deze bevindingen duiden erop dat de verdeling van middelen binnen het huishouden een mogelijke verklaring is voor de daling in scholingsuitkomsten met geboortevolgorde.

The Tinbergen Institute is the Institute for Economic Research, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam and Vrije Universiteit Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. The following books recently appeared in the Tinbergen Institute Research Series:

391. M.I.S.H. MUNANDAR, *Essays on economic integration*.
392. K.G. BERDEN, *On technology, uncertainty and economic growth*.
393. G. VAN DE KUILEN, *The economic measurement of psychological risk attitudes*.
394. E.A. MOOI, *Inter-organizational cooperation, conflict, and change*.
395. A. LLENA NOZAL, *On the dynamics of health, work and socioeconomic status*.
396. P.D.E. DINDO, *Bounded rationality and heterogeneity in economic dynamic models*.
397. D.F. SCHRAGER, *Essays on asset liability modeling*.
398. R. HUANG, *Three essays on the effects of banking regulations*.
399. C.M. VAN MOURIK, *Globalisation and the role of financial accounting information in Japan*.
400. S.M.S.N. MAXIMIANO, *Essays in organizational economics*.
401. W. JANSSENS, *Social capital and cooperation: An impact evaluation of a women's empowerment programme in rural India*.
402. J. VAN DER SLUIS, *Successful entrepreneurship and human capital*.
403. S. DOMINGUEZ MARTINEZ, *Decision making with asymmetric information*.
404. H. SUNARTO, *Understanding the role of bank relationships, relationship marketing, and organizational learning in the performance of people's credit bank*.
405. M.Â. DOS REIS PORTELA, *Four essays on education, growth and labour economics*.
406. S.S. FICCO, *Essays on imperfect information-processing in economics*.
407. P.J.P.M. VERSIJP, *Advances in the use of stochastic dominance in asset pricing*.
408. M.R. WILDENBEEST, *Consumer search and oligopolistic pricing: A theoretical and empirical inquiry*.

409. E. GUSTAFSSON-WRIGHT, *Baring the threads: Social capital, vulnerability and the well-being of children in Guatemala.*
410. S. YERGOU-WORKU, *Marriage markets and fertility in South Africa with comparisons to Britain and Sweden.*
411. J.F. SLIJKERMAN, *Financial stability in the EU.*
412. W.A. VAN DEN BERG, *Private equity acquisitions.*
413. Y. CHENG, *Selected topics on nonparametric conditional quantiles and risk theory.*
414. M. DE POOTER, *Modeling and forecasting stock return volatility and the term structure of interest rates.*
415. F. RAVAZZOLO, *Forecasting financial time series using model averaging.*
416. M.J.E. KABKI, *Transnationalism, local development and social security: the functioning of support networks in rural Ghana.*
417. M. POPLAWSKI RIBEIRO, *Fiscal policy under rules and restrictions.*
418. S.W. BISSESSUR, *Earnings, quality and earnings management: the role of accounting accruals.*
419. L. RATNOVSKI, *A Random Walk Down the Lombard Street: Essays on Banking.*
420. R.P. NICOLAI, *Maintenance models for systems subject to measurable deterioration.*
421. R.K. ANDADARI, *Local clusters in global value chains, a case study of wood furniture clusters in Central Java (Indonesia).*
422. V.KARTSEVA, *Designing Controls for Network Organizations: A Value- Based Approach.*
423. J. ARTS, *Essays on New Product Adoption and Diffusion.*
424. A. BABUS, *Essays on Networks: Theory and Applications.*
425. M. VAN DER VOORT, *Modelling Credit Derivatives.*
426. G. GARITA, *Financial Market Liberalization and Economic Growth.*
427. E.BEKKERS, *Essays on Firm Heterogeneity and Quality in International Trade.*
428. H.LEAHU, *Measure-Valued Differentiation for Finite Products of Measures: Theory and Applications.*
429. G. BALTUSSEN, *New Insights into Behavioral Finance.*

430. W. VERMEULEN, *Essays on Housing Supply, Land Use Regulation and Regional Labour Markets.*
431. I.S. BUHAI, *Essays on Labour Markets: Worker-Firm Dynamics, Occupational Segregation and Workplace Conditions.*
432. C. ZHOU, *On Extreme Value Statistics.*
433. M. VAN DER WEL, *Riskfree Rate Dynamics: Information, Trading, and State Space Modeling.*
434. S.M.W. PHILIPPEN, *Come Close and Co-Creat: Proximities in pharmaceutical innovation networks.*
435. A.V.P.B. MONTEIRO, *The Dynamics of Corporate Credit Risk: An Intensity-based Econometric Analysis.*
436. S.T. TRAUTMANN, *Uncertainty in Individual and Social Decisions: Theory and Experiments.*
437. R. LORD, *Efficient pricing algorithms for exotic derivatives.*
438. R.P. WOLTHOFF, *Essays on Simultaneous Search Equilibrium.*
439. Y.-Y. TSENG, *Valuation of travel time reliability in passenger transport.*
440. M.C. NON, *Essays on Consumer Search and Interlocking Directorates.*